An evaluation of a JUMP Math pilot programme in Ireland

Eemer Eivers, Emer Delaney
and Seán Close

Educational Research Centre

Contents

List of Acronyms and Abbreviations

BLM	Blackline Masters
CBU	Confidence Building Unit
CPD	Continuing Professional Development
DPMT	Drumcondra Primary Mathematics Test
ERC	Educational Research Centre
IMPACT	Interactive Methods and Practical Approaches to Communication and Thinking
JUMP	Junior Undiscovered Math Prodigies
MKT	Mathematical Knowledge for Teaching
MKTQ	Mathematical Knowledge for Teaching Questionnaire
MKTQ-S	Mathematical Knowledge for Teaching Questionnaire – Shortened
MQI	Mathematical Quality of Instruction
NCCA	National Council for Curriculum and Assessment
OCUP	Ontario Curriculum Unit Planner
PDST	Professional Development Service for Teachers
PSMC	Primary School Mathematics Curriculum
SD	Standard Deviation
SIGMA-T	Standardised Irish Graded Mathematics Attainment Test
SME	Subject Matter Expert
TIMSS	Trends in International Mathematics and Science Study
WNCP	Western and Northern Canadian Protocol

Acknowledgements

We gratefully acknowledge the financial support of the Department of Education and Skills (DES), Science Foundation Ireland, and Accenture, and the helpful input of the JUMP Math Steering Committee. The Steering Committee members were Ciarán Cannon (TD, Minister of State for Training and Skills 2011-14), Eamon Clavin (DES), Seán Delaney (Marino Institute of Education), Anne Feerick (DES), Pamela Fitzgerald (Change Nation), Aoife Kelly (National Council for Curriculum and Assessment), Bernard Kirk (Galway Education Centre), and Frank Walsh (Athlone Education Centre). The Educational Research Centre (ERC) was represented by Eemer Eivers, Emer Delaney, and Seán Close.

For the professional development provided to participants, we thank Dr John Mighton from JUMP Math, and Ciara O'Donnell, Mia Treacy, and Michael James Malone from the Professional Development Service for Teachers. Bernard Kirk, Frank Walsh, and staff at Galway and Athlone Education Centres organised Continuing Professional Development (CPD) for participating teachers, and acted as helpful points of contact for the various stakeholders.

Thanks are due to those who conducted our observations, including Dr Dolores Corcoran, Margaret Cooney, John Curran, Pat Moore, and Susan Murray. A number of staff at the ERC provided much appreciated support, including Peter Archer (Acting Director), Paula Chute, Aidan Clerkin, Lauren Kavanagh, Mary Lewis, David Millar, Adrian O'Flaherty, Mary Rohan and Hilary Walshe.

Finally, warmest thanks are due to all participating teachers, pupils, and principals, without whom the evaluation could not have been carried out.

Chapter 1
Introduction

Junior Undiscovered Math Prodigies, or JUMP Math as it is more commonly known, is a Canadian-designed programme intended to help children succeed at, and enjoy, learning mathematics. During the academic year 2013-14, a pilot evaluation of JUMP Math was carried out in Ireland. The evaluation focused on pupils in Third class, and used a sample of Irish primary schools from the catchment areas of two Education Centres, Galway and Athlone.

This report describes the evaluation process and analyses the results. The main functions of the report are to:

- outline the genesis of the project, the key characteristics of JUMP Math, the research design, and the support provided to participating schools (all outlined in this chapter).
- summarise the assessment methods used (Chapter 2).
- analyse the JUMP Math materials provided to participating classes, comparing their content with the objectives of the Irish Primary School Mathematics Curriculum (PSMC) and with mathematics textbooks and materials currently used in Irish primary schools (Chapter 3).
- evaluate the fidelity with which participating schools implemented the programme, and describe the mathematical quality of programme-based instruction, using the outcomes of classroom observations conducted at two points during the school year (Chapter 4).
- establish the effects, if any, of JUMP Math on pupils' and teachers' attitudes to mathematics, drawing on the responses to questionnaires and interviews (Chapters 5 and 6).
- establish the effects, if any, of JUMP Math on pupils' mathematical achievement and teachers' mathematical knowledge for teaching, using the results of baseline and end-of-year standardised tests (Chapter 7).
- summarise the findings and the conclusions that can be drawn from the evaluation (Chapter 8).

Genesis of the project

Change Nation is a social innovation platform which promotes joint ventures by social innovators and national policy makers. Dr John Mighton was among the social innovators who attended its Dublin summit in March 2012, presenting an outline of JUMP Math. Subsequently, with the support of the (then) Minister of State for Training and Skills, Ciarán Cannon, a Steering Committee was set up to examine the feasibility of implementing JUMP Math in Ireland. Sufficient funding was secured from public (Department of Education and Skills, Science Foundation Ireland) and private (Accenture) sources to allow the programme to be implemented in a small number of schools and to fund an evaluation of that implementation. For practical reasons, the programme was limited to schools within the catchment area of two Education Centres – Galway and Athlone – and to a single grade level (Third class).

Key features of JUMP Math

This section provides some information about the background of the JUMP Math programme (hereafter referred to as JUMP), the key principles underpinning it, and an indication of how JUMP is manifested in practice in classrooms.

Background

The JUMP programme was developed by Dr John Mighton in 1998 to tutor individual pupils who were struggling with mathematics. It was subsequently tested and adapted in classroom settings in Canada. JUMP became a registered charity in 2002, and its printed resources for teachers and pupils became available from 2003. Some of the ideas on which JUMP is based are outlined in Mighton's books, *The Myth of Ability* (2003) and *The End of Ignorance* (2007). Put simply, the programme claims to enable a higher standard of mathematics teaching, to improve the performance of all pupils, and to be highly cost-effective. Further information on the programme itself can be found at https://jumpmath.org, including programme philosophy, some research outcomes related to JUMP, and sample teacher and pupil materials. A support network for participating teachers is also available at http://jumpmathteachers.org.

Principles

The principles underpinning JUMP are described in the Introduction section of the JUMP teacher manuals (Mighton, Sabourin & Klebanov, 2010). The seven core principles are:

- confidence-building.
- guided practice.
- guided discovery.
- continuous assessment.
- rigorously scaffolded instruction.
- mental maths.
- deep conceptual understanding.

In addition, JUMP is guided by one overarching principle, which is that *all* pupils remain fully engaged with lessons.

The following sections elaborate on what each principle entails. It is, however, worth noting that the meaning ascribed to some of the terms used may not always tally with that ascribed in the general mathematics education literature. For example, if teaching methodologies were considered on a continuum from total 'Didacticism' to 'Pure Discovery', many might place JUMP's methodologies well away from the Discovery end of the continuum, yet guided discovery is one of JUMP's seven core principles. This is because JUMP's definition of guided discovery is somewhat different to that used generally. Thus, readers should note that the next sections draw heavily on Mighton et al.'s (2010) description of JUMP, and are intended purely as description, not critique.

Confidence-building

The JUMP programme is based on the belief that every child can be good at mathematics, and that a confident and interested child is more likely to learn in an efficient manner. Thus, it aims to decrease pupils' anxiety about mathematics, and to minimise the differences pupils perceive between their own ability and that of their classmates. To do this, Mighton et al. (2010)

advocate “raising the bar” in small increments and providing pupils with prompt, positive feedback.

The JUMP method advocates that most or all of the first fortnight of the school year is spent using a “Confidence Building Unit” (CBU), designed to promote pupil confidence at whole-class level. The CBUs are grade-specific, and the Third class version covers addition and subtraction of fractions (Mighton, 2013a and 2013b). The content is not part of the Irish curriculum for Third class (or the Ontario curriculum for Third grade), and is closer to the content that might be expected at Fourth class or higher. One of the reasons fractions is the topic chosen for the CBU is that it requires little text and is mainly composed of mathematical symbols (numerals zero to nine and some operation signs). The CBU is not intended as an in-depth review of fractions. As the teacher manual to the CBU indicates, “the goal of the unit is not to teach students to fully understand the connection between the symbolic operations with fractions and the concrete models that underlie those operations … but to allow children to experience complete mastery in a rich and interesting abstract game” (Mighton, 2013a, p. 6). Thus, the aim is to show all pupils, but particularly weaker pupils, that they can complete difficult-looking challenges.

Guided practice

Mighton et al. (2010) argue that repetition and practice are key to learning, and that component skills must be repeatedly practised before “the big picture” can be understood. They believe that “working memory” is limited and can have difficulty processing a large amount of new information. With this in mind, the JUMP *Assessment and Practice Workbooks* (Mighton, Sabourin, & Klebanov, 2009) given to pupils are intended – in conjunction with JUMP lesson plans – to provide opportunity for guided practice and repetition.

Guided discovery

Mighton et al. (2010) define guided discovery as allowing pupils to take moderate risks by exploring tasks that are within their grasp, yet have an element of inquiry-based learning. They suggest that discovery should form a part (but not the entirety) of any well-planned lesson, with independent work balanced by explicit hints and instructions. More recently, Mighton (2013c) has elaborated further on the drawbacks he sees to a “pure discovery” approach. In particular, he suggested that the use of too many illustrations and concrete materials could prove distracting, and that discovery learning could place too heavy a burden on working memory.

Continuous assessment

The JUMP approach requires teachers to make regular use of continuous assessment activities. The teacher manuals show how to reduce learning into discrete steps and how to assess each step. In addition to “mini-quizzes” and tasks following the introduction of each new skill or concept, JUMP advocates a monthly cumulative review, using a selection of related workbook questions copied onto a single sheet. Teachers are also warned that the purpose of continuous assessment is not ranking, but to allow pupils to show their knowledge and teachers to “differentiate instruction with small individual interventions” (Mighton et al., 2010, p. 5).

Separately, Mighton (2007) has described his ideal approach to assessment as involving “visible” and “invisible” assessment methods. Visible assessments are marks awarded to pupils for particular tasks or tests. There would be only two grade levels: “A” and, for pupils who had done particularly well, “A+”. Teachers would proceed to the next topic only once all pupils had achieved one of these grades. Invisible assessment would consist of something similar to a diagnostic profile for each pupil. This would be passed between teachers as pupils progressed through grade levels, and shared with parents but not with pupils.

From an Irish perspective, it is worth comparing how the PSMC and JUMP differ in relation to continuous assessment. The PSMC also places a strong emphasis on assessment, but specifies the importance of *formative* as well as summative assessment and suggests tools for producing formative data: for example, pupil portfolios and curriculum profiles (National Council for Curriculum and Assessment & Department of Education and Science, 1999). In contrast, JUMP methods – at least, as outlined in the teacher manuals – seem designed to produce largely what the PSMC would define as summative data (although Mighton et al. [2010, p. 5] do use the term "formative" to describe JUMP assessment methods). While Mighton's *theoretical* perspective recognises the importance of formative ("invisible") data, as defined in the PSMC, tools such as pupil portfolios are not discussed in the JUMP manuals and are therefore unlikely to be used by teachers adhering strictly to JUMP methodologies. However, individual teachers can adapt standard JUMP assessment methods to yield formative and diagnostic data.

Rigorously scaffolded instruction

The JUMP teacher manuals argue that pupils can only retain a limited amount of new information or steps at any given time, meaning that all instruction should be broken down into sequential, scaffolded steps, with strong direct instructional guidance from the teacher.

The lesson plans in the manuals contain suggestions for how to break procedures into simpler stages. For example, when teaching mixed fractions, a teacher might break the procedure of drawing these into two steps. First, the pupils could learn to draw the number of wholes present, before learning to draw the fractional part.

Mental maths

In the JUMP teacher manuals, Mighton et al. (2010) argue that "mental math is the foundation for all further study in mathematics" (p. 6). Familiarity with mathematical patterns, through consistent practice, is believed to facilitate pupils' understanding of how numbers interact. In turn, they can then calculate quickly without having to recall number facts.

Deep conceptual understanding

Mighton et al. (2010) say that JUMP proposes the simultaneous teaching of symbolic and concrete understanding of key mathematical concepts, using a variety of approaches. An aim of JUMP is that teachers can see mathematical "big ideas" even in small steps, while learning how to relate the small steps to wider contexts and procedures.

Mighton (2003, p. 58) explains that "even in the most basic units, students are expected to explain how operations work and generalize rules to deal with new cases by themselves". While he believes that an approach that teaches rules before concepts can have major benefits for the confidence of struggling pupils, the ultimate aim of JUMP is to create conditions in which concepts can be taught before rules, and pupils can discover mathematical principles on their own.

Keeping all pupils engaged

JUMP aims to keep all pupils engaged by providing challenges that grow incrementally harder in the form of bonus questions – designed to look demanding (e.g., by using larger numbers than usual), but which rely on skills and concepts that pupils have already mastered. Bonus questions are intended to keep stronger pupils engaged without teacher intervention, enabling attention to be diverted to any pupils who are struggling. Examples of bonus questions are included in all pupil workbooks and teacher manuals (which also suggest strategies for creating further bonus questions).

As well as maintaining individual pupil engagement, JUMP seeks to create collective, whole-class enthusiasm. Pupils moving through materials at faster speeds do not move to the next topic, but instead complete bonus questions on the same topic until the whole class can proceed. The idea is that knowledge gaps among pupils can thus be closed, and that pupils will experience success as a group. Mighton (2007) describes the anticipated enthusiasm as a form of "collective effervescence", following Émile Durkheim.

JUMP also aims to make pupils excited about mathematics by pointing to "the beauty of mathematics as a symbolic language connected to the real world" (Mighton et al., 2010, p. 2). Thus, in the mental mathematics strategies, emphasis is placed on patterns (e.g., the fact that the sum of digits in any two-digit multiple of nine is always nine).

JUMP in practice

The preceding sections outlined the key principles underpinning JUMP. For the purposes of the evaluation, it was important to operationalise these principles – that is, to define how adherence to abstract principles might be manifested in classroom practice, and in a manner that could be measured. However, it should also be noted that many of the more aspirational JUMP principles broadly align with what is commonly considered good classroom practice. For example, as noted earlier, continuous assessment is a key element of the Irish PSMC. As such, a crude measure of how regularly teachers assessed pupils might not distinguish between teachers following JUMP or PSMC principles.

That caveat aside, it seems likely that mathematics teaching that adhered closely to the JUMP methodology would include the following features:

- use of the JUMP "Confidence Building Unit" (CBU) at the start of the year.
- frequent use of "Mental Math" strategies, as described in the teacher manuals.
- provision of highly scaffolded instruction, following the detailed lesson plans described in the teacher manuals.
- use of bonus questions as a means of offering differentiated teaching.
- progression through topics at whole-class level.
- drawing attention to "beautiful" aspects of mathematics and making real-world links.
- frequent assignment and correction of mini-quizzes and cumulative reviews.
- large amounts of repetition and practice of mathematical concepts.
- large amounts of teacher-led questioning and explanation.

In addition, as JUMP pays considerable attention to building pupil confidence, and retaining high levels of engagement throughout the school year, it would seem likely that pupils experiencing successful implementation of the programme should develop significantly higher levels of mathematical self-confidence, and report greater interest in, and enjoyment of, mathematics.

Research Design

In June 2013, Galway Education Centre and Athlone Education Centre invited interested schools to apply to participate in a small-scale evaluation of the JUMP programme in an Irish context. The invitation provided some background information on JUMP and an outline of methods to be used for the evaluation. Schools were advised that in addition to the outcomes of standardised tests, the evaluation would include classroom observations of mathematics lessons, and interviews with teachers and pupils. The invitation indicated that selected schools would receive:

- free training in the methodologies underpinning JUMP.
- free JUMP classroom materials (teacher manuals and pupil workbooks) for a school year.

In total, 23 schools applied. The Steering Committee had planned that 12 schools would be selected for the JUMP programme, with the remaining schools used as a control group against which performance could be compared. Non-selected schools were to be offered the possibility of participation in a JUMP programme in the 2014/15 school year. Schools were to be selected with reference to a geographical balance between areas covered by the two Education Centres.

Having recruited schools, an invitation to tender for the research evaluation was issued in July 2013, and the work subsequently awarded to the Educational Research Centre (ERC). ERC staff proposed a revised design, outlined in the next section.

Revised research design

The main change to the study design was an attempt to address a) a probable Hawthorne Effect[1] for selected schools, and b) a demotivating effect for non-selected schools. As noted, non-selected schools were to be offered the possibility of participation in a JUMP programme in the 2014/15 school year. However, this might have had the unintended consequence of a de-emphasis on mathematics in those schools during the 2013/14 school year. To counteract this and any Hawthorne Effect, the ERC proposed a second comparison intervention as part of the evaluation. The Steering Committee agreed to the proposed change.

A number of maths programmes were examined (by staff at the ERC and in the Education Centres) as potential comparison interventions. The Professional Development Service for Teachers' (PDST) new programme – IMPACT Maths (Interactive Methods and Practical Approaches to Communication and Thinking) – was selected. There were three main reasons for choosing IMPACT. First, like JUMP, it was directed at all pupils in a class rather than targeting only weaker or higher achieving pupils. Second, in contrast to JUMP, it was developed to align very closely with the content and pedagogical framework of the Irish PSMC. Third, in contrast to JUMP, it adopts a social constructivist approach, with emphases on pupils sharing and discussing mathematics, and following broad-based learning trajectories. JUMP, on the other hand, breaks mathematics into highly specific skills and concepts which are taught through teacher-dominated questioning and explanation, followed by graded individual pupil practice.

[1] The Hawthorne Effect is the phenomenon whereby people tend to improve or modify an aspect of their behaviour once they are aware that they are being studied. Commonly known methods to reduce such an effect include the use of placebos, assessing more than one intervention, or double blind studies.

Although IMPACT materials were not available for all strands (only covering Shape and Space, and part of Number), the programme's principles and methods are intended to be generalisable across strands. This meant that it was suitable as an overall comparison group, and addressed the Hawthorne Effect. However, the fact that no materials were available for the Measures, Data, and Algebra strands should be borne in mind when interpreting programme effects in later chapters.

While the final evaluation design included a comparison of two programmes, and both were given equal importance in all communications from the ERC to schools, it is possible that some teachers selected for the IMPACT group felt they were assigned to a lesser programme, or were not allocated their first choice. This is not due to programme characteristics, but to the fact that the initial invitation referred to JUMP only. Thus, the possible perception of IMPACT as a lesser programme may mean that a Hawthorne Effect could have operated differently in the two groups.

Participants

Of the 23 schools that applied to participate, two withdrew in September 2013. One was replaced, giving a final sample of 22 schools, of which 12 were in the Galway Education Centre catchment area and 10 in the Athlone Education Centre catchment area. As most schools had a single Third class group, the total number of participating class groups was 27 (13 in Galway and 14 in Athlone).[2] All schools but one were mixed-sex. In 14 schools, the Third class group(s) were single-grade, while eight schools had multi-grade classes (the most common combination being Third and Fourth classes). The sample was reasonably balanced by rural/urban location (15 versus seven schools, respectively) – urban being defined as a school situated in a town with a population in excess of 5,000. In practice, all but one of the participating urban schools were in towns with population in excess of 20,000. Although fewer schools were in urban than rural areas, the larger average size of the urban schools meant that pupils were evenly split by location (280 in urban schools and 289 in rural schools).

Method of assigning schools to programmes

The small numbers of participating schools meant random allocation to programmes was unlikely to create two balanced, matched groups. Thus, schools were assigned to sample groups using an iterative process, to maximise comparability across the following key variables:

- prior achievement on standardised tests.
- number of schools, classes and pupils.
- geographical location (urban/rural).

Schools supplied the ERC with anonymised achievement data from standardised tests of mathematics achievement administered at the end of the 2012/13 academic year. These data were used to ensure that there were minimal pre-existing differences in mathematics achievement between the JUMP and IMPACT groups. Some schools had used the Drumcondra Primary Mathematics Test (DPMT) while others had used the Standardised Irish Graded Mathematics Attainment Test (SIGMA-T). As scores on different standardised tests could not be assumed to be directly equivalent, schools were assigned to programmes in a way that minimised difference not only on overall mean achievement, but also on DPMT and SIGMA-T scores.

[2] Initially, 28 classes were listed as participating. However, one JUMP school which had applied with two participating class groups decided to combine these for mathematics lessons during the pilot year.

Alongside prior achievement, numbers, and urban/rural location, two other factors were also taken into account – neighbour effects and balance between the Education Centres. As it was possible that schools located close to each other might share aspects of their programme that they found useful, any neighbouring schools were assigned to the same condition. In addition, an effort was made to achieve similar proportions of JUMP/IMPACT schools in the Athlone and Galway areas.

Characteristics of final samples

Table 1.1 summarises characteristics of schools and pupils assigned to each of the interventions. Twelve schools (13 class groups containing 295 Third class pupils) were allocated to JUMP, while 10 schools (14 class groups containing 274 Third class pupils) were allocated to IMPACT Maths.[3] The JUMP group comprised eight rural schools (176 pupils) and four urban schools (119 pupils), while the IMPACT group comprised seven rural schools (113 pupils) and three urban schools (161 pupils). Seven schools in each group had single-grade classes, while five JUMP and three IMPACT schools had multi-grade classes.

Mean prior achievement scores for the two groups were within a point of each other (108.5 for JUMP pupils, 109.4 for IMPACT pupils). Both groups contained a mixture of pupils who had been administered the DPMT and the SIGMA-T. As well as a close match on overall mean achievement, the means for the DPMT and SIGMA-T were closely matched across groups. In sum, both the JUMP and IMPACT groups of schools and pupils were closely matched on all key variables, including initial pupil mathematical achievement.

Table 1.1: Summary characteristics of JUMP and IMPACT school, class, and pupil participation

		JUMP	IMPACT	Total
Schools *N*		12	10	22
Classes *N*		13	14	27
Pupils *N*		295	274	569
Location (schools *N*)	Rural	8	7	15
	Urban	4	3	7
Education Centre (schools *N*)	Athlone	5	5	10
	Galway	7	5	12
Class structure (schools *N*)	Single-grade	7	7	14
	Multi-grade	5	3	8
Mean achievement score		108.5	109.4	108.9[4]

Nature of support provided to schools

Teachers in both programmes received tailored Continuing Professional Development (CPD) and additional materials. The nature and extent of support to be provided to the JUMP programme participants was based on advice from JUMP personnel, and had been established prior to the involvement of the ERC. Therefore, support for teachers in the IMPACT programme was tailored to mirror that offered to JUMP participants.

[3] Initially, schools were slightly more evenly split. However, one teacher from a school assigned to IMPACT subsequently informed the ERC that she was familiar with, and used, JUMP methodologies. Her school was therefore re-assigned to the JUMP group.
[4] Means may not match exactly, due to rounding.

Continuing Professional Development (CPD)

Over the course of the evaluation, teachers were invited to three CPD sessions related to their assigned programme. The first of these took place on Saturday, September 7th 2013. JUMP training was delivered by Dr John Mighton, while IMPACT training was delivered by three representatives of the PDST. Each group of teachers was welcomed to the session by the then Minister of State, Ciarán Cannon, and the Director of Athlone Education Centre, Frank Walsh, and each session lasted approximately six hours. As far as was possible, the two programmes were accorded equal status.

The JUMP training included illustrations of how the programme's principles might be put into practice – for instance, how to scaffold a lesson on parts and wholes, how to devise bonus questions to motivate pupils, and how to point out number patterns so as to excite pupils' mathematical curiosity. A video of Dr Mighton using the JUMP methodology to teach a lesson on perimeter was shown to demonstrate some of these points. There was little specific reference to the JUMP teacher manuals. The IMPACT training included discussions of the programme's philosophy of "teaching in the concrete and pictorial for understanding in the abstract" and techniques such as re-voicing. Unlike the JUMP training, it incorporated small-group discussions, group exercises using concrete materials, and references to specific pages of the (two) IMPACT manuals available at that time. Thus, in certain ways, the manner of CPD delivery mirrored the different philosophies of the two programmes. JUMP CPD was very much led by the "teacher", targeted at the whole group, and involved little shared learning. In contrast, IMPACT was more "student"-led, with a somewhat social constructivist approach to CPD.

The sessions were attended by at least one teacher from 69% of JUMP classes and 71% of IMPACT classes. In most cases, the class teacher attended, but in a few cases, it was the resource teacher, or both class and resource teachers. The slightly low attendance rate can largely be attributed to the short notice given to teachers (as little as six days in some cases) for training delivered on a Saturday. A delay in accessing an element of the project funding led to a delay in establishing Dr Mighton's availability (around which *both* sessions were scheduled), all of which contributed to the short notice given to teachers.

For those who could not attend, recordings of the sessions were placed on the Galway Education Centre website. Follow-up in December 2013 revealed that only two non-attending teachers from each group had watched the recordings. Thus, two classes in each programme were taught for a full term by teachers who had neither attended the initial training session, nor accessed the recorded version of it.

The next two CPD sessions for each programme were in webinar form. The first of these took place on November 25th (JUMP) and 26th (IMPACT). Teachers were invited to log in either at Galway Education Centre or from their home or school. Dr Mighton facilitated the JUMP webinar, while a PDST representative facilitated the IMPACT one. The JUMP webinar dealt predominantly with questions raised by participating teachers, while the IMPACT webinar dealt with the strand for which materials had most recently been developed, i.e., Shape and Space. Across both groups, eight classes had at least one teacher in attendance (information is unavailable for one IMPACT teacher). Again, low attendance can be at least partly attributed to the short notice given to participating teachers. Short notice was largely due to the difficulty of scheduling around mid-term, Dr Mighton's availability, generally, and his availability in a suitable time zone, specifically.

Combining data on attendance or online viewing of the September session with participation in the November webinars revealed that one JUMP class and two IMPACT classes were taught by teachers who had not received *any* related CPD up to the end of 2013.

The second set of webinars had a similar format, and took place on February 10th (JUMP) and 11th (IMPACT). At least one teacher attended from 69% of JUMP classes and 86% of IMPACT classes. (A few JUMP teachers who tried to log in could not do so, due to technical problems). For JUMP teachers, the webinar involved another question and answer session, with Dr Mighton providing examples of how teachers might tackle specific problems they were having. For IMPACT teachers, the webinar involved further suggestions on teaching Shape and Space in accordance with the principles of the programme. Teacher participation in CPD is summarised in Table 1.2.

Table 1.2: Percentage of classes from which teachers participated in CPD, by programme

	JUMP (N=13)	IMPACT (N=14)
Teacher(s) attended live CPD	69	71
Teacher(s) watched video of live CPD	15	15
Teacher(s) took part in first webinar	62	62
Teacher(s) took part in second webinar	69	86

Materials

Each teacher was supplied with materials appropriate to their assigned programme.

JUMP materials

The JUMP materials are analysed in detail in Chapter 3. Broadly, they consisted of teacher manuals for levels 3.1 and 3.2; pupil workbooks, also levels 3.1 and 3.2; and teacher and pupil versions of the supplementary "Confidence Building Unit", level C. These materials are intended to cover the complete Ontario and Western and Northern Canadian Protocol mathematics curricula at Grade 3 (broadly equivalent to Third class in Ireland). The manuals and pupil materials were very detailed, totalling 731 pages for teachers (including answer keys and Blackline Masters materials), and 349 pages for pupils. Following the first webinar, some JUMP teachers requested Level 4 materials also. Pupil workbooks 4.1 and 4.2 (but not the matching teacher manuals) were delivered to all JUMP classes in January 2014.

IMPACT materials

The IMPACT materials consisted of teacher manuals for three strands/strand units from the Irish PSMC: Fractions; Place Value, Percentages and Decimals; and Shape and Space. However, the PDST representatives emphasised during CPD that the intention was to provide *examples* of how the programme's principles could be used in particular strand units, but that teachers could themselves apply the principles to other areas of the curriculum. In correspondence with the PSMC, the Number manuals dealt with teaching the content to all levels from First through to Sixth class, while the Shape and Space manual dealt with levels from Junior Infants to Sixth class. In total, the manuals were 430 pages in length, with 96 pages specifically relevant to teaching Third class. No pupil materials accompanied the IMPACT teacher manuals, which recommended the use of readily available classroom items in maths lessons.

Chapter 2
Assessment methods

The study used a wide variety of measures to evaluate the pilot JUMP programme. First, JUMP materials were analysed to establish how they compared with the Irish PSMC, and with materials currently used in Irish schools. Second, classroom observations were conducted at two points during the year to assess the fidelity with which the programmes (both JUMP and IMPACT) were implemented, and to observe any notable differences in mathematical quality of instruction between the JUMP and IMPACT groups. Third, all pupils' and teachers' attitudes to mathematics were measured using start- and end-of-year questionnaires. Fourth, interviews were conducted with teachers and with a small selection of pupils. Fifth, teacher mathematical knowledge for teaching (MKT) was assessed pre- and post-intervention. Sixth, standardised tests were administered at the start and end of the year, to assess pupil mathematical achievement pre- and post-intervention.

This chapter describes the nature of each of the assessment methods used, while the outcomes of the assessments are described in later chapters.

JUMP materials and the Irish curriculum

To examine the degree of alignment with the Irish Primary School Mathematics Curriculum, JUMP materials were compared first to the intended curriculum (using the PSMC objectives) and second to the curriculum as commonly implemented (using several widely used mathematics textbooks). The comparative procedures used are outlined next, while findings are reported in Chapter 3.

PSMC content objectives

The PSMC content objectives for Third class were listed, and all level 3 JUMP teacher manuals and pupil workbooks were analysed to assess the extent to which each objective was addressed. For each objective, a rating was applied to indicate whether it was fully addressed, partly addressed, or not addressed by the JUMP materials. The analysis was jointly conducted by two staff members in the ERC, one of whom was a mathematics Subject Matter Expert (SME).

Irish textbooks

As there may be discrepancies between the intended curriculum and the curriculum as reflected through textbooks and related materials, JUMP resources were also compared with three commercially available Irish mathematics textbooks for Third class. The three textbooks chosen were all in use in participating schools prior to the evaluation, with one textbook in particular used in a large majority of participating classrooms. The materials do not represent a finite set of all materials encountered by pupils in Third class. Other textbooks may be in use in a small number of classes, nationally. However, between them, the textbooks selected provide a reasonably representative picture of how the Irish PSMC is typically addressed in Irish mathematics textbooks, and, in turn, in many mathematics lessons.

Conceptual analysis of mathematics textbooks and related materials is still a developing research area (Remillard, 2005), particularly where cross-cultural comparison is concerned. An instrument designed by Charalambous, Delaney, Hsu and Mesa (2010) specifically for cross-cultural textbook comparison was identified as the most suitable template from which to develop a means of textbook comparison for this evaluation. Charalambous et al.'s instrument had not only been used to compare textbook series from three different countries, but had been used successfully to identify characteristics that distinguished *Irish* mathematics textbooks from those of other countries. The instrument includes two forms of analysis: "horizontal", providing a holistic picture of each textbook, and "vertical", focusing on a single content area (the addition and subtraction of fractions, in the case of their own detailed analyses). Within the vertical analysis, worked examples and pupil tasks can be coded on several dimensions.

Charalambous et al.'s instrument was adapted for use in the current study, in consultation with Dr Seán Delaney (one of its developers). First, since addition and subtraction of fractions do not feature in the PSMC at Third class level, vertical analysis was performed on another element of the Fractions strand unit (equal parts in relation to fractions). Second, the analysis of worked examples was omitted, as there were almost no examples of these in the selected samples of either JUMP or the Irish textbook. Third, the codes used to classify pupil tasks in terms of their *potential cognitive demands* and *response required* were adapted. Charalambous et al. had reported difficulty in distinguishing between the two *potential cognitive demand* codes of "procedures with connections" and "procedures without connections". Therefore, a revised set of codes was developed to avoid this problem, drawing on categories used in Irish assessment tools such as the National Assessments of Mathematics (Eivers et al., 2010) and the DPMT. Of the four codes used by Charalambous et al. for *response required*, only one ("single answers") was present in the Irish textbooks they examined. Initial analysis of the JUMP and Irish comparison materials suggested that there would be a somewhat similar pattern in both. Therefore, the four codes were simplified to two: "closed response" (corresponding with "single answers"), and "open/extended response" (which might include an explanation, a justification, or an answer along with the mathematical sentence used to reach it – thus covering all three of Charalambous et al.'s other codes). Fourth, since much of JUMP's content is communicated via the teacher manual, the textbook analysis was extended to include teacher manuals as well as pupil workbooks, and the coding system modified to include teacher prompts as well as pupil tasks. A copy of the adapted instrument is included in Appendix A.

Classroom observations

Observations took place at two points during the school year. The first set of observations took place as soon as possible after the first webinar, which was held at the end of November 2013. Most were conducted in December 2013, with three taking place in January 2014. The second set of observations took place in May 2014, immediately before the second round of pupil testing began. All but one of the participating class groups were observed twice. The non-observed class was one of three Third classes in a school. As it proved impractical to conduct three sequential observations and associated interviews in a single day, (the same) one of the classes was not observed on either occasion.

For each set of observations, slightly more than half of lessons were recorded and coded at a later date, while the remainder were observed and coded "live". Each recorded lesson was observed and coded by two subject matter experts (SMEs), while each live observation was completed by a former primary school principal. As the recorded observations allowed for more in-depth analyses, some of the data collected in September 2013 were used to ensure that the classes selected for the recorded condition represented a broad spectrum of class level mathematics achievement and teacher MKT scores.

Eight classes from each programme were assigned to the recorded observation format for both December and May observations. However, as one JUMP school had combined two Third classes for maths lessons, data were collected for only 15 class groups (seven from JUMP and eight from IMPACT). The remaining classes (six from JUMP and five from IMPACT) were assigned to the live observation format.

Observation instruments

Three different observation instruments were used. First, for all observations, an observation schedule designed specifically for the purposes of the study was used. Next, for the recorded observations only, a lesson report and the Mathematical Quality of Instruction (MQI) instrument (Learning Mathematics for Teaching Project, 2011) were used. The instruments are explained next, while the findings of all observation analyses are described in Chapter 4.

Observation schedule

In Chapter 1 (*JUMP in practice*), some likely characteristics of mathematics lessons adhering to JUMP methodology were identified. These characteristics (and some expected characteristics of IMPACT lessons) were used to inform the development of a tailored observation schedule (Appendix B). The schedule's key function was to gauge the extent to which teachers adhered to the methodologies of their respective programmes, while also providing more general information about classroom dynamics.

The schedule was used for both live and recorded observations. In a live setting, it was used by former primary school principals, while in the recorded lessons, it was used by one of the SMEs. Observers, all of whom had been trained in correct use of the schedule, used it to describe grouping methods, instructional practices, pupil behaviour, classroom climate, and types of materials used. In addition, based on what they had observed and on their understanding of the programmes, observers used their professional judgement to offer global ratings of teacher adherence to JUMP or IMPACT.

Lesson report

For each recorded lesson, both SMEs produced a brief outline of the lesson, including a description of the lesson's main activities, the time spent on each, and the materials used. The outline provided an overall picture of classroom practice, and was used to supplement findings on programme adherence and mathematical quality of instruction. See Appendix C for the report template.

Mathematical Quality of Instruction (MQI)

The MQI instrument (Learning Mathematics for Teaching Project, 2011) was used by both SMEs to rate recorded observations as a number of its dimensions corresponded closely to claims made for the JUMP programme. For instance, the MQI measures "Explanations", "Mathematical language", and "Remediation of student errors and difficulties", all areas specifically targeted by JUMP teaching resources. It provides overall estimates of Mathematical Quality of Instruction and Mathematical Knowledge for Teaching, drawing upon summed ratings on four broad dimensions:

- richness of mathematics (including linking, explanations, developing generalisations, and fluency of mathematical language).
- working with students and mathematics (including error remediation and responding to student mathematical questions).

- errors and imprecision (including teacher error in language or notation, clarity of presentation of content).
- student participation in meaning-making and reasoning (including pupil explanations, reasoning, engaging with the content at a high cognitive level.)

The full MQI rates lessons in seven-minute segments. However, as the tailored observation schedule provided information on the *content* of lesson segments and qualitative lesson reports provided information on the *structure* of lessons, the MQI was used in a truncated form (see Appendix D). Rather than a segmented approach, it was used to generate a holistic assessment of the mathematical quality of lessons.

Combining data from the observation schedules and MQI, it was possible to measure adherence to JUMP methods and to establish how adherence to methods was associated with *quality* of instruction.

Questionnaires

To gather contextual information, questionnaires were administered to teachers and pupils at the start and end of the school year. Although containing slightly different content, both Teacher Questionnaires asked about materials and teaching practices used in maths lessons, the amount of time allocated to maths lessons, and the extent of teachers' confidence in teaching aspects of maths. The initial questionnaire also asked for some demographic information, while the second questionnaire asked some questions about their experience of and views on the programme to which they had been assigned. The first questionnaire was completed in September 2013, and the second was completed prior to the end-of-year achievement test.

Pupil Questionnaires asked similar questions on both occasions, including questions about pupils' attitudes to school in general, and mathematics in particular. They also asked about homework and pupils' usual activities during maths lessons. At both the start- and end-of-year testing, the achievement test was first completed, then, after a short break, the Pupil Questionnaire was completed, all under the supervision of an ERC test administrator. Copies of all questionnaires (pupil and teacher) are included in Appendix E.

Interviews

All teachers and a sample of pupils were interviewed (using tailored interview schedules) after classroom observations – i.e., in December 2013 or January 2014, and again in May 2014. Appendix F contains the schedules used for teachers and pupils.

Teacher interview

The teacher interview was designed to gauge teachers' views of their own adherence to JUMP or IMPACT, and their attitudes to their programme at the time of interview. The second set of interviews was slightly more detailed than the first, and included teachers reflecting on which aspects of their programme they had used during the year, which (if any) they believed to be particularly effective, and their general views on the programme evaluation. On both occasions, one interview was conducted for each class observed. For five class groups (three JUMP and two IMPACT), interview responses reflected the combined views of a class teacher and a resource teacher.

Pupil interview

The pupil interview was designed to gauge pupil attitudes to mathematics. As with teachers, the interview schedule was adapted slightly for the second administration: questions were rephrased to provide maximum clarity to respondents, and questions which had yielded minimal responses were deleted. In each class observed, three pupils were selected by their teacher as representative of pupils with low, medium, and high mathematical achievement. The three were then interviewed as a group by the ERC observer. The composition of the group of three was decided by the class teacher and was not the same in each set of interviews.

Teacher mathematical knowledge

The Mathematical Knowledge for Teaching Questionnaire (MKTQ) is, as the name suggests, a measure of teacher knowledge of mathematical concepts and ability to apply that knowledge in the context of mathematics lessons. Based on work by Hill, Schilling and Ball (2004) and Delaney, Ball, Hill, Schilling and Zopf (2008), it was normed in Ireland on a sample of 500 teachers (Delaney, 2012). Delaney's data were analysed to see if it were possible to reduce the length of the questionnaire without losing discriminatory power. As a result of the analyses, 42 of the original 84 items were retained to form a shortened version (MKTQ-S), highly correlated (r=.96) with the full-length MKTQ.

The MKTQ-S was completed by teachers in the study on two occasions (September 2013 and May/June 2014). In September, it was typically completed just prior to the start of initial CPD for JUMP/IMPACT. Teachers unable to attend that session completed the MKTQ-S while their pupils were being assessed for the first time. The MKTQ-S was again completed by teachers at the end of the school year, while their pupils were being assessed.

Pupil achievement

The Drumcondra Primary Mathematics Test (DPMT) is a set of standardised mathematics achievement tests developed specifically for group administration in Irish primary schools. The tests are divided into six Levels, most of which have parallel forms. The tests are designed for pupils at the end of First class through to the end of Sixth class. For Third class, Level 2 is used for autumn testing and Level 3 is used for end-of-year testing. Originally developed in 1997, the DPMT was completely revised in 2005. The revised tests (also known as the DPMT-R) reflect the aims and structure of the 1999 Primary School Mathematics Curriculum. Test results from over 16,000 pupils were used to develop norms for each Level of the test. Test items are evenly split between multiple-choice and short-answer format.

The start-of-year test window for administration of the DPMTs was between September 9th and 20th 2013. The end-of-year test window was between May 13th and June 6th 2014, although almost all schools tested between May 20th – 30th. On both occasions, schools agreed a suitable morning for the administration of the tests, and all tests were administered by external test administrators (either ERC staff or qualified primary school teachers). The test administrators returned all test materials to the ERC where they were scored by ERC staff.

Response rates

Table 2.1 shows response rates for all instruments in both September 2013 and May/June 2014. In September, 546 of 569 pupils (96%) completed the DPMT and the Pupil Questionnaire. Of the 23 non-participants, 16 pupils were absent and seven pupils' parents refused permission for them to take part. Response rates were also very high for the MKTQ-S and Teacher Questionnaire. Of the 28 class teachers (including job-sharing posts), all but one completed the MKTQ-S (96%), while all completed the Teacher Questionnaire.[1] Thus, response rates for the initial data collection phase were excellent.

Similarly, response rates for May/June 2014 were high, with 536 pupils (94%) completing the DPMT and the Pupil Questionnaire. However, teacher response rates decreased somewhat, with 22 class teachers (79%) completing the MKTQ-S, and 25 (89%) completing the Teacher Questionnaire.

Table 2.1: Response rates for pupil- and teacher-level instruments, September 2013 and May/June 2014

	September 2013			May/June 2014		
	N	Completed	Response rate	N	Completed	Response rate
DPMT	569	546	96.0%	569	536	94.2%
Pupil Questionnaire	569	546	96.0%	569	536	94.2%
MKTQ-S	28	27	96.4%	28	22	78.6%
Teacher Questionnaire	28	28	100.0%	28	25	89.3%

Summary

A variety of methods were used to evaluate the piloting of JUMP methods in a small number of Third classes. JUMP materials were compared with the Irish PSMC, and with materials for Third class currently used in Irish schools. Classroom observations were conducted at two points during the year to assess how the programmes (both JUMP and IMPACT) were being implemented. Pupils' and teachers' attitudes to mathematics were measured using start- and end-of-year questionnaires, and interviews were conducted with teachers and with a small selection of pupils. Teacher mathematical knowledge for teaching was assessed pre- and post-intervention, and standardised tests were administered at the start and end of the year, to assess pupil mathematical achievement pre- and post-intervention.

[1] As class teachers were the target respondents, rates reflect the percentage of class teachers who provided data. However, five resource teachers who indicated that they engaged in significant amounts of team teaching also volunteered data. Their responses were combined with those of the class teachers to produce a mean result for each class group.

Chapter 3
Analysis of materials

The materials provided as part of the JUMP programme, including teacher manuals and pupil workbooks, are analysed in this chapter. First, an overview of JUMP materials is provided. Second, their content is analysed to see how fully they address the objectives of the Irish PSMC. Third, JUMP materials are compared with three sets of commercially available Irish mathematics materials for Third class pupils. The materials are broadly compared in terms of their relative weighting of the curriculum strands, and emphasis on computational practice. Following this, the treatment of a single topic is analysed in detail, in JUMP and in the set of Irish materials most commonly used by participating teachers.

The analysis of materials takes place in a contextual vacuum, examining what might be learned *if* teachers were to follow the manuals with no deviations, and *if* pupils were to solve every exercise in their workbooks (Mesa, 2004). Thus, while the present chapter considers the JUMP *intended* curriculum, later chapters consider factors mediating the creation of an *enacted* curriculum (Remillard, 2005). As such, analyses here should be read in conjunction with findings on teachers' adherence, and attitudes, to programme materials (Chapters 4 and 6).

Overview of JUMP materials

As indicated in Chapter 1, the materials provided to participating JUMP teachers consisted of:

- teacher manuals, levels 3.1 and 3.2.
- pupil workbooks, levels 3.1 and 3.2 (and, from January 2014, 4.1 and 4.2).
- teacher and pupil versions of the supplementary "Confidence Building Unit", level C.

These are described in more detail in the sections that follow.

Teacher manuals

The JUMP teacher manuals are extensive, containing:

- an introduction to the principles and methods of JUMP Math (15 pages).
- a detailed sample template of a "problem-solving lesson" on perimeter (five pages).
- a section on mental mathematics skills, exercises, and assessment (19 pages).
- lesson plans, organised in five strands: Patterns and Algebra; Number Sense; Measurement; Probability and Data Management; and Geometry. All strands are represented in both manuals, in the above order. (471 pages).
- Blackline Masters materials, including charts and games that could be photocopied or used as templates from which to create materials. Suggested uses are given in the lesson plans. (98 pages).
- answer keys for pupil workbooks 3.1 and 3.2 (42 pages).
- tests (with answer keys) for each of the five strands (55 pages).
- tables summarising the correspondence between JUMP materials and the Grade 3 objectives of two Canadian curricula (26 pages).

Thus, lesson plans comprise the bulk (64%) of the teacher manuals. A typical lesson plan begins with a summary of goals, prior knowledge required, and useful vocabulary. Next, it provides specific instructions as to what the teacher should ask, tell, draw, etc. Plans usually end with Activities and/or Extensions (often including "bonus questions"), through which a topic can be further developed.

Appendix G shows some sample JUMP lesson plans, addressing the topics of Place Value, Writing and Reading Number Words, Writing Numbers, and Representation with Base Ten Materials, with accompanying Blackline Masters and pupil materials.

Pupil workbooks

Pupil workbooks are called Assessment and Practice workbooks. For Grade 3, the workbooks comprise:

- a brief note for teachers and parents on how to use JUMP Math (one identical page in each volume).
- worksheets, organised in the same strands used by the teacher manuals (349 pages).

The pupil workbooks for Grade 4 are similarly arranged (349 pages of worksheets).

All workbooks are printed in black and white, although the covers feature brightly-coloured images. The exercises are rarely accompanied by graphical displays other than representations. Sample extracts are included in Appendix G, addressing the same topics (Place Value, Writing Numbers, and Base Ten) as the sample teacher lesson plans in that appendix.

Confidence-Building Unit (CBU)

The CBU (Level C) is intended for use at the start of the school year, for a maximum of two weeks. As outlined in Chapter 1, it deals with the addition, subtraction, and reduction of fractions, but does not aim to teach the topic in *conceptual* depth. Rather, it aims to promote pupils' confidence by assuring them that they can master *procedures* usually tackled by pupils at more advanced grade levels. The CBU comprises:

- a teacher manual (for Levels C and D combined, but with the recommendation that only Level C be used with Grade 3 pupils). The manual explains the purpose of the CBU, and provides lesson plans, sample homework, tests, answer keys, and an appendix on teaching basic operations. (62 pages in total, of which 42 are relevant to Level C).
- a pupil workbook ("Fractions Challenge: Level C"). This comprises worksheets on basic operations, adding and subtracting fractions, reducing fractions, and naming mixed and improper fractions. (30 pages).

While the JUMP programme generally aims to balance procedural and conceptual learning, guided practice and guided discovery, the CBU represents procedural learning and guided practice in distilled form. The lesson plans in the CBU teacher manual feature even more "steps" than those in the main teacher manuals, but rarely suggest explaining the rationale behind steps. (The teachers' introduction to the CBU observes that fractions can be taught in depth using the main JUMP teacher manual and pupil workbook, but that the purpose of this unit is "more to build confidence, harness attention and motivate children to learn their number facts than to teach fractions completely" [Mighton, 2013a, p. 10]).

Of course, many commercial textbooks and teacher manuals also do not include the rationale behind activities. However, with a typical textbook series, teachers use the core material to develop their own lesson plans (presumably including reference to rationale). In contrast, JUMP presents an apparently fully-formed lesson plan, so the exclusion of conceptual explanation in the CBU becomes more important.

Two other relatively unusual features of the CBU are that teachers and pupils are presented with only one concrete model (the circle/pie) and one algorithmic approach to any problem, and in the pupil workbook, there are *no* graphical displays other than representations (with the exception of one page featuring drawings of hands, intended to aid pupils in skip counting). Thus, of all the JUMP Math materials, the CBU seemed likely to be perceived by participating teachers as least similar to typical Irish materials. As shall be seen in Chapter 6, teachers in just under half the JUMP classes indicated that they did not use the CBU, while those who did use it had varied views on its usefulness.

JUMP Structure

JUMP has been developed to support the Ontario Curriculum Unit Planner (OCUP), and consequently shares some of its structural characteristics. JUMP and OCUP – like the PSMC – are grouped around five strands. OCUP contains 65 specific objectives, and either two or three “Overall Expectations” for each strand unit (these are not counted in the 65 objectives as they summarise the more specific aims). Each specific OCUP objective is the target of at least one JUMP lesson unit. However, 10 objectives (15%) are addressed fully in the teacher’s lesson plan, but not in the pupil workbooks. Examples include objectives relating to temperature, the comparison of objects by units of mass or capacity, and the comparison of angles using concrete materials and pictorial representations. In other cases, multiple lesson plans and workbook activities are directed towards a single OCUP objective – e.g., nine lesson plan units address the objective to “represent, compare, and order whole numbers to 100, using a variety of tools”.

Thus, while JUMP is designed to cover the Ontario curriculum in full, not all content is in the pupil workbooks, meaning coverage is contingent on teachers adhering to the content of their manual. Also, some objectives receive considerably more attention than others. However, this may reflect the complexity of some objectives relative to others. For example, the single objective related to representing, comparing, and ordering whole numbers requires that teachers explain a number of concepts and demonstrate operations in a variety of ways. Other objectives, such as the objective to “use a reference tool to identify right angles and to describe angles as greater than, equal to, or less than a right angle”, may be taught within a single lesson unit.

The five strands in JUMP (Number Sense, Measurement, Geometry, Patterns and Algebra, and Probability and Data Management) not only closely mirror those of the OCUP, but also broadly correspond to the five strands of the Irish PSMC (Number, Measures, Shape and Space, Algebra, and Data) (Table 3.1).

Table 3.1: Strands in the Ontario curriculum (OCUP), JUMP materials, and the Irish primary curriculum (PSMC)

OCUP	JUMP	Irish PSMC
Number Sense and Numeration	Number Sense	Number
Measurement	Measurement	Measures
Geometry and Spatial Sense	Geometry	Shape and Space
Patterning and Algebra	Patterns and Algebra	Algebra
Data Management and Probability	Probability and Data Management	Data

JUMP materials and the PSMC

As noted in the preceding section, although JUMP was originally based on the strands and objectives of the OCUP, there is broad similarity between the five strands underpinning both JUMP and the PSMC. This section looks in more detail at the relative importance accorded to each strand in JUMP and the PSMC, and examines JUMP materials to see how well, if at all, each of the PSMC content objectives for Third class is addressed. Where we indicate that an objective is "addressed", this simply means that most or all content relevant to an objective is dealt with by JUMP materials. As with any set of materials, coverage of key content is a necessary prerequisite for a learner successfully meeting an objective, but does not, by itself, ensure that the learner will *meet* the objective. While the PSMC also lists overarching skills that a learner should master, these are not included in the present analysis as it is not possible to quantify the extent to which they are addressed in any set of materials.

Strand emphasis in JUMP and PSMC

To measure how emphasis is divided between strands in JUMP and the PSMC, the percentages of JUMP lesson units and PSMC objectives per strand were quantified. However, as this is a slightly crude indicator of the importance accorded to strands, this was followed by a more detailed review of *how* JUMP and the PSMC deal with each strand. JUMP for Grade 3 contains 230 lesson units, while the PSMC has 70 specific curriculum objectives relating to Third class. As shown in Table 3.2, a little less than half (42%) of the JUMP lesson units fall under Number Sense. This is more than double the number of lesson units allocated to Geometry (19%), and considerably more than allocated to Patterns and Algebra (16%) and Measurement (14%). Only one in ten JUMP lesson units addresses topics related to Probability and Data Management.

In contrast, 36% of PSMC objectives relate to the Number strand and 24% to each of Measures and Shape and Space. Although coverage of Data is broadly in line with JUMP's coverage of Probability and Data Management, only 7% percent of PSMC objectives relate to Algebra. At first glance, therefore, JUMP places more emphasis than the PSMC on two strands (Number and Algebra), less emphasis on Shape and Space, and considerably less emphasis on Measures.[1]

Table 3.2: Percentage of JUMP lesson units and PSMC objectives per strand (using their own classification systems)

JUMP (N=230)		PSMC (N=70)	
Strand	% lesson units	Strand	% objectives
Number Sense	41.7	Number	35.7
Measurement	13.9	Measures	24.3
Geometry	19.1	Shape & Space	24.3
Patterns & Algebra	15.7	Algebra	7.1
Probability & Data Management	9.6	Data	8.6

However, a single objective may be addressed over several lessons, or a single lesson may address several objectives. Further, while the general content of the JUMP and PSMC strands match quite well, there are differences in classification. Therefore, a more detailed

[1] For simplicity, the names of the PSMC strands are also used to refer to their JUMP equivalents.

analysis is needed to understand if and how strands are treated differently in JUMP and the PSMC.

Number

The strand Number Sense composes a larger percentage of JUMP lesson units than the percentage of PSMC objectives falling under the Number strand. However, JUMP's Number strand includes the treatment of money, which falls under the Measures strand in the PSMC. Excluding lesson units related to money brings a slightly closer alignment between the emphasis on Number in JUMP and in the PSMC.

Also, certain Number objectives require multiple lessons due to the mental mathematics and operational skills involved. For example, the two PSMC objectives "explore and identify place value in whole numbers, 0-999" and "read, write and order three-digit numbers" are addressed across *nine* JUMP lesson units. That JUMP also goes well beyond 999 and addresses numbers in the thousands and ten thousands may account for part of the greater emphasis in JUMP lesson plans on Number. However, it is likely the greater emphasis largely reflects the fact that while it is possible to summarise the two PSMC objectives relatively succinctly, the actual content, the skills and procedures involved, requires a considerable amount of class time. Put simply, some objectives take longer to teach than others. This, coupled with the "extra" money strand units, suggests that JUMP and the PSMC match quite closely in terms of the relative emphasis placed on the Number strand.

Measures

The Measurement strand does not feature very prominently in JUMP lesson units. However, as noted in the preceding section, this is partially attributable to money being classified under Number Sense, and two units on time being classified under Patterns and Algebra. Within the Measurement strand, JUMP allocates only one lesson unit to weight and one to capacity, whereas two curriculum objectives are related to each in the PSMC. In contrast, JUMP covers temperature and perimeter, neither of which feature in the PSMC for Third class. In sum, JUMP appears to place relatively less emphasis than the PSMC on Measures, but this is mainly attributable to money appearing under the Number strand.

Shape and Space

JUMP units address all Shape and Space topics dealt with by the PSMC, and some additional material not contained in the PSMC (congruency, as in Euclidean geometry; use of grids, as in coordinate geometry; and in-depth study of flips, slides and turns, as in transformational geometry). While the PSMC requires pupils to recognise an angle in terms of a rotation, the Ontario curriculum requires pupils to identify flips, slides, and turns, through investigation using concrete materials and physical motion, and to name flips, slides, and turns as reflections, translations, and rotations. Reflecting the heavy emphasis in OCUP, JUMP dedicates 12 lesson units to flips, slides, and turns.

Despite this, the percentage of JUMP lesson units addressing Shape and Space is slightly low with respect to PSMC objectives, partly because the PSMC seeks a more in-depth study of 2-D and 3-D shapes than JUMP provides. In JUMP, 2-D shapes that are polygons and 3-D shapes that are prisms and pyramids are considered, but the associated activities promoted by the PSMC are not all included. For example, JUMP does not require pupils to tessellate 2-D shapes. Also, the PSMC objective on parallel, horizontal and vertical lines is not fully reflected in JUMP lesson units (although prior knowledge of horizontal and vertical lines is assumed in some units). Thus, it seems that JUMP and the PSMC emphasise different aspects *within* the strand of Shape and Space. Also, while JUMP covers a large number of topics, it does not cover all of them in great depth.

Algebra

The Patterns and Algebra strand in JUMP receives proportionally more attention (16% versus 7%) than the broadly equivalent Algebra strand in the PSMC, largely reflecting the greater emphasis on the strand found in the Ontario curriculum. The main topics covered are similar (i.e., number sequences and number sentences/equations). However, JUMP also includes a wide variety of *types* of sequences – e.g., pupils are challenged to identify, describe and extend number patterns involving addition, subtraction, and multiplication, represented on a number line, a calendar, or a hundreds chart. JUMP also includes four units dealing with patterns in the times tables (as an aid to mental mathematics). In contrast, the PSMC requires pupils to explore, extend and describe sequences, but does not specify what *kinds* of sequences.

Classification by two criteria is included in the Patterns and Algebra strand of JUMP for Third grade, but is covered at Second class in the PSMC. More generally, JUMP has five Patterns and Algebra lesson units related to classification that the PSMC would categorise under Shape and Space. It also has two units on time patterns that would be more likely to fall under Measures on the PSMC. If classification- and time-related content is excluded, the difference in emphasis between the PSMC and JUMP reduces, but only to 7% of objectives versus 13% of lesson units. Thus, in the case of Algebra, it seems likely that JUMP materials provide a little more depth and breadth of coverage than indicated by the PSMC.

Data

Data was the strand least emphasised in both JUMP and the PSMC, although some JUMP content related to data went beyond the aims of the PSMC. For example, JUMP devotes a lesson unit to the understanding of fairness in games and its relationship to probability, something not addressed at Third class in the PSMC. Also, JUMP includes varieties of data representation – such as Venn diagrams – not mentioned explicitly in the PSMC (or, indeed, in the OCUP). Those caveats aside, JUMP and the PSMC seemed generally well matched on the Data strand.

Overall balance

The five strands of the JUMP materials correspond closely with those of the Irish PSMC. Although some topics appeared under different strands in JUMP and the PSMC, there was a high level of correspondence in content. Table 3.2 showed the percent of JUMP and PSMC content under each of their respective strands. However, this did not take into account situations where there is a close match on content covered in both, but where the strand classification differs (as in the case of money).

Applying the PSMC strand structure to JUMP lesson units, units related to money or time would move from Number Sense, and Patterns and Algebra (respectively), to Measurement/Measures. Units related to classification would move from Patterns and Algebra to Geometry/Shape and Space. Doing this reveals that JUMP and PSMC content is very similar (Table 3.3). JUMP's apparent heavier emphasis on Number disappears, as does the apparent lesser emphasis on Measures. It is only in the treatment of content in the Algebra strand that notable differences remain between the emphases found in JUMP and in the PSMC.

Table 3.3: Percentage of JUMP lesson plans and PSMC objectives that fall under the PSMC strands (using PSMC definition of strand content)

	JUMP (N=230)	PSMC (N=70)
Number	34.8	35.7
Algebra	12.6	7.1
Shape & Space	21.3	24.3
Measures	21.7	24.3
Data	9.6	8.6

PSMC objectives not addressed

The JUMP materials were analysed by two researchers to assess the extent to which they addressed each of the 70 objectives of the PSMC for Third class. An objective was deemed fully addressed if the JUMP teacher manual and/or pupil workbook could be used without major adaptations to facilitate the targeted outcome. For instance, the objective that pupils should be able to "read, write and order three-digit numbers" was deemed fully addressed, as at least six JUMP lesson units dealt with it in detail. While some of the terminology was slightly different ("three hundred [and] one"), this was not perceived to interfere with the objective being fully and clearly addressed.

An objective was deemed partly addressed if the JUMP teacher manual and/or pupil workbook could be used to achieve some elements of the targeted outcome, but would need significant adaptation to achieve the outcome in full. For instance, the objective to "multiply a one-digit or two-digit number by 0-10" was rated as only partly addressed. This was because JUMP dealt with multiplication of one-digit numbers, but had limited coverage of multiplication of two-digit numbers (the few exceptions involved doubling, and multiplying larger numbers by 10).

An objective was deemed not addressed if the JUMP teacher manual and/or pupil workbook did not provide any obvious starting point from which to proceed towards the targeted outcome.

Although the review process was somewhat subjective, both reviewers independently agreed that of the 70 PSMC objectives, 63 were fully addressed by JUMP materials, six were partly addressed, and only one not addressed at all. Full details of the review are contained in Appendix H, which maps each of the 70 objectives onto JUMP lesson unit(s), and identifies differences in approach, if any, between JUMP and the PSMC. Table 3.4 draws on Appendix H to present information about the relatively few objectives that JUMP did not fully address.

As can be seen, the only objective not even partly addressed was the aim that pupils develop an understanding of the relationship between fractions and division. However, this objective is in practice rarely addressed by Irish-produced materials either. Other strand units in which objectives were only partly addressed included Number: Operations (two objectives), Algebra: Number sentences, and Shape and Space: Lines and angles. Also, as might reasonably be expected, two objectives from the Measures strand relating to money were considered to be only partly addressed, due to JUMP's use of Canadian currency rather than euro.

Of course, the fact that, theoretically, JUMP and PSMC content are broadly aligned does not necessarily mean that teachers found they aligned well in practice. Therefore, the preceding analyses should be read in conjunction with data from teachers' interviews (Chapter 6). In particular, while we rated JUMP money sections as "partially addressing" the relevant PSMC objectives, some JUMP teachers felt the pupil materials related to money were unusable.

Thus, in the case of money and perhaps a small number of other objectives, while the theoretical alignment was good, the practical implementation (i.e., unmodified use of JUMP pupil materials to teach the topic) was difficult.

It is also notable that a significant minority of JUMP teachers in the first set of interviews considered the "pitch" of the materials wrong for Third class (i.e., either too high, too low, or a mixture of too high and too low). A few – inaccurately – believed that Canadian Grade 3 was equivalent to Irish Fourth class and that, therefore, the materials were not grade-appropriate. However, by the second set of interviews, only one JUMP teacher raised pitch as an issue.

Table 3.4: PSMC objectives which were partly addressed or not addressed by the JUMP materials

Strand: *strand unit*	PSMC Objective	Relevant JUMP unit(s)	Gaps in JUMP
Number: *Operations*	Multiply a one-digit or two-digit number by 0-10	NS3-36- 39.	Limited instances of multiplying two-digit numbers (e.g., doubling, multiplying larger numbers by 10).
Number: *Operations*	Divide a one-digit or two-digit number by a one-digit number with and without remainders	NS3-62- 63; NS3-66.	Does not require that work be recorded using the division algorithm (as PSMC does).
Number: *Fractions*	Develop an understanding of the relationship between fractions and division	n/a **Not addressed**	Only one unit (NS3-85) deals with fractions greater than one. No explicit link with division was made.
Algebra: *Number sentences*	Translate an addition or subtraction number sentence with a frame into a word problem (frame not in initial position)	PA3-33 and 35; NS3-88-91.	Focus on translating word problems to number sentences, rather than vice versa.
Shape & Space: *Lines & angles*	Identify, describe and classify vertical, horizontal and parallel lines	G3-11 – 14, especially G-12.	Deals with horizontal and vertical lines in the context of symmetry, but does not deal with parallel lines.
Measures: *Money – euro*	Rename amounts of euro or cents and record using symbols and decimal point	NS3-42-47; NS3-70-74.	Money section referenced Canadian currency, not euro.
Measures: *Money - euro*	Solve and complete one-step problems and tasks involving the addition and subtraction of money	NS3-48; NS3-75-76.	Money section referenced Canadian currency, not euro.

JUMP materials and Irish textbooks

The preceding sections outlined the structure of JUMP materials and their relationship to the PSMC. However, daily mathematics instruction in Irish primary classrooms is usually based around a textbook, not the curriculum (Eivers et al., 2010). As the textbook is the medium through which the primary mathematics curriculum is experienced, a brief comparison of JUMP materials and three commercially available Irish textbooks was carried out. These three were chosen as they were the only ones that teachers in the study mentioned using.

The JUMP pupil workbooks and the Irish materials were analysed in their entirety, to see what proportions of each set of materials focused on the various strands. In addition, a more detailed "vertical" analysis (similar to Charalambous et al., 2010) examined how a single mathematical concept was dealt with in different materials. The concept selected for review was that of *equal parts in relation to fractions.* One of the three sets of Irish materials was chosen for the more detailed comparison with JUMP, on the basis that it was the textbook series most widely found in participating classrooms prior to the evaluation, and it was from one of the two Irish textbook series analysed in the Charalambous et al. study.

Overview: JUMP pupil workbooks and three Irish textbooks

JUMP pupil workbooks were compared with three commercially available pupil textbooks for Third class. These materials are hereafter referred to as Textbooks A, B and C, where Textbook A was the book used in a large majority of classrooms prior to the evaluation.

Earlier in this chapter, number of JUMP lesson units was used as a measure of relative strand emphasis. As lesson units do not have direct equivalents in pupil materials in Irish textbook series, page counts were used as an indicator of relative coverage. All three Irish textbooks include revision sections for pupils, which were included in page counts. Revision in JUMP materials tends to be covered in teacher materials, and could not therefore be included in the analysis of pupil materials. However, as the relative topic emphasis in revision sections largely mirrored the relative topic emphasis in the main parts of pupil materials, the exclusion of JUMP revision sections does not substantively alter the balance between strands. In a related vein, supplementary pupil materials for some of the Irish textbooks were not included in the analyses, as not all pupils use them. However, content of supplementary materials largely mirrored the relative emphases found in the "parent" textbook.

In the JUMP pupil workbooks, 50% of pages related to the PSMC Number and Algebra strands (Table 3.5).[2] This is slightly higher than the 43% of PSMC objectives related to Number and Algebra but nonetheless lower than the percentages in all three sets of Irish textbooks – 53% in Textbook C, 61% in Textbook B, and 65% in the widely used Textbook A.

At 21%, level of coverage of Measures in JUMP was very similar to levels in Irish materials (23% in Textbook C, and 20% in both Textbook B and Textbook A). Generally, the amount of cover given to Measures in both JUMP and Irish textbooks is quite close to the 24% of PSMC objectives devoted to Measures.

In contrast, Shape and Space, which accounts for 24% of PSMC objectives, received proportionally more coverage in the JUMP workbooks (20% of pages) than in any of the Irish materials (17% in Textbook C, 14% in Textbook B, and only 8% in the most widely used textbook, Textbook A). JUMP pupil materials were also similar to the PSMC in terms of coverage given to Data (9% in both). However, Data received relatively little coverage in the three Irish textbooks (7% in Textbook C, 6% in Textbook A, and 5% in Textbook B).

Table 3.5: Percentages of pages in JUMP pupil workbooks and three Irish pupil textbooks that cover each PSMC strand

% of PSMC objectives (N=70)		% of pages			
		JUMP (N=349)	Textbook A (N=174)	Textbook B (N=172)	Textbook C (N=156)
Number & Algebra	42.8	49.6	65.2	61.2	52.9
Shape & Space	24.3	20.1	8.3	13.5	16.5
Measures	24.3	21.2	20.1	20.5	23.4
Data	8.6	9.2	6.3	4.8	7.2

[2] The comparison takes account of strand classification differences discussed earlier in this chapter – e.g., Money was counted under Measures rather than Number and Algebra in JUMP.

Comparing the JUMP workbooks with the textbooks normally used in Irish classrooms, it appears that JUMP gives less emphasis to Number and Algebra, and more emphasis to Shape and Space. It would not be reasonable to expect a perfect match between the percentage of PSMC objectives per strand and the percentages of pages dealing with content in that strand. As noted earlier, for example, certain Number objectives are quite complex, and require multiple lessons. A slight "overemphasis" on Number in all pupil materials might be expected, while still anticipating a broad alignment between percentage of objectives and percentages of pages. However, the focus on Number in some Irish textbooks (particularly the widely used Textbook A) is more than a slight overemphasis, and contributes to relatively limited coverage of the Shape and Space and/or Data strands. In contrast, strand emphasis in JUMP materials is more closely aligned with PSMC objectives than in any of the Irish textbooks considered.

Aside from strand coverage, the percentage of pages containing purely computational questions – i.e., questions involving varied and repetitive practice of counting, addition, subtraction, multiplication, and division skills, in isolation from any practical context or verbal representation – was examined. Despite JUMP's emphasis on repeated procedures, only 16% of pages in the JUMP workbooks dealt solely with computation – a lower percentage than was found in *any* of the Irish materials. Coverage in Textbook C and Textbook B was reasonably similar to JUMP, with 19% and 18% of their pages relating to computation, respectively. In contrast, the very popular Textbook A devoted 29% of its pages to computation.

A closer look: "equal parts" in JUMP and Textbook A

For the more detailed "vertical analysis", Charalambous et al.'s framework was considered a useful starting point for developing the review method, as it had previously been used for cross-cultural comparison of textbook series, including comparing Irish mathematics textbooks with those from other countries. However, their target concept for vertical analysis (*addition and subtraction of fractions*) did not feature at Third class. Therefore, the analysis was performed on another element of the Fractions strand unit to retain as much similarity in method as possible. "Equal parts" was chosen as a specific focus because the Fractions sections of both JUMP and Textbook A materials began by considering this concept. As JUMP teacher manuals contain materials that do not appear in the pupil workbooks, the textbook analysis was extended to include teacher manuals as well as pupil workbooks, and the coding system modified to include teacher prompts as well as pupil tasks.

JUMP materials for analysis included units NS3-78 ("Equal Parts"), NS3-79 ("Models of Fractions"), and NS3-81 ("Equal Parts of a Set"), in both the pupil workbook and teacher manual (3.2). Textbook A materials for analysis included two pages from the first section on fractions in the pupil textbook, and two pages from a supplementary pupil workbook,[3] as well as associated instructions from the teacher manual. The comparability of the mathematical content of these sections in JUMP and Textbook A was confirmed by an SME prior to the development of a coding scheme.

Pupil materials

The tasks presented in pupil materials were analysed on two dimensions: the **potential cognitive demands** they made of pupils, and the **type of response** they required. Following

[3] While the supplementary pupil workbook was excluded from the previous analysis of strand and computational content, it was included here as the focus was on specific *items*. It was thought that the items in the supplementary workbook might differ somewhat from those in the main textbook, as the two books are intended to complement one another (even if in practice they are not always used together).

Charalambous et al. (2010), a task was defined as any question asked within the exercises/problems, even if it was not individually numbered. The selected extract from the JUMP pupil workbook contained 73 tasks, while the extracts from Textbook A's pupil textbook and supplementary workbook contained a total of 82 tasks.

Potential cognitive demands of tasks were categorised by Charalambous et al. using four codes, based on the Task Analysis Guide (Stein, Smith, Henningson, & Silver, 2000). These were: memorisation; procedures without connections; procedures with connections; and doing mathematics. However, Charalambous et al. reported particular difficulties in distinguishing between the categories of procedures *without* connections and procedures *with* connections. Therefore, the coding system was adapted so that procedures requiring connections were grouped with other tasks that involved higher-order cognitive demands.

The three types of demands were therefore:

- **recall of facts, terms, or concepts** (corresponding roughly with Charalambous et al.'s memorisation code).
- **implementation of procedures** (procedures without connections).
- **reasoning, connecting, or problem-solving** (broadly similar to Charalambous et al.'s "doing mathematics" code, but also including procedures with connections).

As an added advantage, the revised task structure aligned reasonably well with the classifications used in assessment instruments used as part of the National Assessments (Eivers et al., 2010).

Most of the tasks in JUMP (74%), but even more in Textbook A (85%), were coded as demanding recall (Table 3.6). However, the remaining 26% of JUMP tasks were all coded as requiring the higher-order skills of reasoning, connecting, or problem-solving. In contrast, the remaining 15% of Textbook A tasks all required implementation of procedures.

Table 3.6: Percentages of tasks in JUMP and Textbook A pupil materials coded under three categories of potential cognitive demands

	JUMP (N=73)	Textbook A (N=82)
Recall of facts, terms, or concepts	74.0	85.4
Implementation of procedures	0.0	14.6
Reasoning, connecting, or problem-solving	26.0	0.0

Tasks were also categorised by the type of response they required. As described in Chapter 2, Charalambous et al.'s codes were reduced to two for the present evaluation:

- **closed response** (corresponding with "answer only" code)
- **extended/open response** (which could include an explanation and/or a justification).

Most of the JUMP tasks (79%) and *all* the Textbook A tasks were coded as requiring a closed response[4] (Table 3.7). The 21% of JUMP tasks requiring an open or extended response were largely composed of tasks demanding reasoning, connecting, or problem-solving.

[4] This mirrors Charalambous et al.'s (2010) finding that Textbook A's (Fifth class level) treatment of addition and subtraction of fractions included only closed response options.

Table 3.7: Percentages of tasks in JUMP and Textbook A coded under two categories of response required

	JUMP (N=73)	Textbook A (N=82)
Closed response	79.5	100.0
Extended/open response	20.5	0.0

Teacher materials

Teacher materials were not included in the framework developed by Charalambous et al. However, it was essential to include them in this analysis since the JUMP programme relies heavily on its teacher manuals both to cover content and to direct teaching strategies. Therefore, a coding system was developed to account for the instructions given in the sample extracts of the JUMP and Textbook A teacher manuals.

At a structural level, differences between the two sets of teacher manuals appeared significantly more pronounced than differences between the pupil materials. As described earlier, the JUMP teacher manual contains very detailed lesson plan units, including instructions to teachers to ask specific questions, draw specific representations, etc. (Three such lesson plans were included in the extract sampled). The Textbook A manual, however, offers a much more general guide to teaching the topic. For the "Fractions" strand unit, it includes the PSMC objectives, a summary of useful terminology, and a list of things to keep in mind (e.g., that it is important to develop a solid understanding of tenths before decimals can be introduced). After this preamble to the strand unit as a whole, there are notes to guide the teacher through each page of the pupil textbook. Notes to the relevant two pages of the pupil textbook were included. The teacher manual for Textbook A includes appendices which can be photocopied and used as pupils' work pages. One of these is relevant to the topic of equal parts and fractions. Therefore, the guide to using this appendix was also included in the sampled extract.

To compare the contrasting approaches of the JUMP and Textbook A teacher manuals, a system was developed to code "teacher prompts". A "teacher prompt" was defined as an instruction or piece of advice given to a teacher. It could be brief or lengthy, specific or vague, provided that it dealt with a single element of the teacher's expected activity. Thus, examples include:

- "Explain that a fraction has a top and a bottom number." – JUMP teacher manual
- "Finding half and quarters of sets is presented in a manner which will support the pupils in the transition from hands-on materials to symbolic representation. These concepts have already been covered in previous lessons." – Textbook A teacher manual.

There were 98 such prompts in the sample JUMP extract, but only 12 in the sample Textbook A extract. This discrepancy underlines the different *functions* of the two manuals: JUMP intends to provide very specific instruction to teachers, while Textbook A intends to offer a more generalised guide, and expects teachers to draw more upon their own resources.

Teacher prompts were coded on two dimensions. First, a "Yes/No" code captured whether the information in a prompt (e.g., a script, a representation, or an instruction) **could be directly copied** by the teacher without further adaptation or preparation. For JUMP, 94 of the 98 prompts (95.9%) contained information that could be directly copied, compared to nine of the 12 Textbook A prompts. Second, prompts were coded based on what the teacher was being prompted to do (Table 3.8).

Categories included:

- verbally **eliciting** information: an answer only, *or* an explanation of an answer/process).
- **explaining** something: a fact, procedure, or step, *or* a deep conceptual reason or mathematical generalisation.
- **drawing or writing** a mathematical representation.
- **initiating** activity: use of structured manipulatives, *or* use of everyday materials, *or* pupil discussion, *or* pupils drawing or writing.
- explicitly **making links**: with other mathematical concepts/procedures, *or* with other school subjects/everyday life.
- mentally **anticipating**: probable pupil error/confusion, *or* connections between present and past/future work, *or* an opportunity for differentiation among pupils.

Table 3.8: Percentages of teacher prompts in JUMP and Textbook A manuals coded under six broad categories of suggested activity, and 14 subcategories

		JUMP (N=98)	Textbook A (N=12)
Verbally elicit	Answer only	33.7	8.3
	Explanation of answer/process	4.1	0.0
Explain	Fact, procedure, or step in procedure	13.3	0.0
	Deep conceptual reason/generalisation	1.0	0.0
Draw or write	Representation	19.4	25.0
Initiate activity	Use of structured manipulatives	2.0	8.3
	Use of everyday materials	2.0	8.3
	Pupil discussion	2.0	8.3
	Pupils drawing or writing	14.3	25.0
Make link	With other maths concepts/procedures	1.0	0.0
	With other subjects/everyday situations	2.0	0.0
Anticipate	Probable pupil error/confusion	3.1	0.0
	Connections with past/future maths learning	1.0	16.7
	Opportunity for differentiation among pupils	1.0	0.0

Unsurprisingly, given the large disparity in total number of prompts, JUMP prompts fell into considerably more categories than did Textbook A prompts. That aside, there were some similarities between the two sets of materials in terms of relative emphasis. Both JUMP and Textbook A frequently prompted teachers to draw or write a representation, or to initiate pupil activity around drawing or writing. However, JUMP placed a heavy emphasis (34% of prompts) on teachers verbally eliciting simple answers from pupils, whereas only one (8%) Textbook A teacher prompt did so. More generally, 50% of JUMP prompts fell into the categories of verbally eliciting and explaining information, while 50% of the Textbook A prompts involved the teacher either drawing or writing themselves, or initiating pupils to draw or write. However, given that there were only 12 Textbook A prompts versus 98 for JUMP, the conclusions that can be drawn from analysis of teacher materials are limited.

Summary

Generally, the *mathematical content* of the JUMP materials is a good match for that of the PSMC. Of the 70 PSMC content objectives, 63 were fully addressed by JUMP materials, six were partly addressed, and only one not addressed at all. While some topics would require cultural adaptation for use in an Irish context (e.g., money), overall JUMP and PSMC content is very similar and it is only on the Algebra strand that notable differences in emphasis can be found.

Differences between JUMP materials and pupil materials normally used in Irish classrooms were more pronounced. The JUMP pupil workbooks were *more* closely aligned with the PSMC, in terms of relative emphasis placed on Number and Algebra, Shape and Space, and Data, than any of the three Irish textbooks considered. The JUMP workbooks also placed less emphasis on isolated computation than did any of the Irish books.

Textbook A differed most from both the PSMC and JUMP, yet was the textbook most commonly used by teachers participating in the evaluation. Comparing it with JUMP materials revealed superficial differences in the use of colour and illustrations, but broad similarities in the demands made of pupils. In the more detailed analysis of how "Equal Parts" was dealt with, most JUMP questions sought recall of facts and concepts, and closed responses, while a small percentage required pupils to engage in higher-order cognitive activities, or to provide extended or open responses. In contrast, *all* Textbook A questions were based on recall and in closed format, while none required higher-order cognitive activities.

However, there were marked differences between the *format and purpose* of JUMP materials and Textbook A. The JUMP teacher manual contains very detailed lesson plans almost entirely composed of scripts, representations, and injunctions that teachers could reproduce directly. In contrast, the Textbook A teacher manual provides a broad, general guide to teaching each topic and to the exercises in the accompanying pupil workbooks. Given the differences in functions of the materials, and the fact that only one topic treatment was examined in detail, considerable caution should be exercised in interpreting differences between the materials.

Finally, as already noted, pupils' experiences in class are not determined by intended curriculum alone, but by a complex interaction of the curriculum materials and the learning environment (Tarr, Reys, Reys, Chávez, Shih & Osterlind, 2008). Eivers et al. (2010) reported a heavy emphasis on the textbook as the basis of mathematical instruction in Irish classrooms, while Charalambous et al. (2010) noted that "the role of textbooks in instruction depends on how students and teachers interact with them in instruction" (p.118). JUMP materials and the PSMC are reasonably well aligned, in theory, but how this manifests in practice must also be considered. Thus, findings in this chapter should be considered alongside Chapters 4 and 6, dealing with programme implementation.

Chapter 4
Classroom observations

As outlined in Chapter 2, classroom observations were conducted for two main reasons. First, they provided an indication of the extent to which programme principles were reflected in teachers' classroom practices. Second, they allowed analyses of differences, if any, between the mathematical quality of instruction experienced by pupils in each programme.

To measure *adherence* to programme principles, a tailored observation schedule was developed (incorporating specific adherence indicators and more global ratings of adherence). To measure the *quality* of mathematical instruction, recorded lessons were rated on a modified version of the Mathematical Quality of Instruction (MQI) instrument (Learning Mathematics for Teaching Project, 2011) and described using qualitative lesson reports.

This chapter is divided into five sections, the first of which explains the types of observation data collected. The second section describes the curriculum areas covered in the lessons observed. The third section describes the extent of adherence to programme methods and principles, and the fourth analyses the quality of mathematical instruction in the lessons observed. The fifth offers some general comments on the observed lessons, based on the overall impressions of observers.

Types of observation data

The content of this chapter is based on 52 mathematics lessons that were observed as part of the evaluation. Twenty-six classes were observed between December 2013 and January 2014, and again in May 2014. On both occasions, trained observers observed lessons "live" in 11 classes, while lessons were recorded for later analysis in 15 classes. In the first set of observations, average lesson duration for both programmes was 44 minutes and ranged between 30 and 60 minutes. In the second set, average duration was again 44 minutes for JUMP lessons (range 23 to 60 minutes), while IMPACT lessons averaged 46 minutes (range 24 to 73 minutes).

Three main types of observation data were collected: those from a tailored observation schedule, ratings from an abridged MQI, and a summary lesson report.

Tailored observation schedule

All 52 lessons (i.e., both live and recorded) were observed using a tailored observation schedule. The activities targeted in the schedule were those expected to produce maximum difference between the two programmes, were teachers to adhere closely to programme principles. Thus, high levels of adherence to JUMP principles and methods were expected to manifest in lessons emphasising high levels of teacher-led instruction, low incidence of group work, frequent use of workbooks, bonus questions, memorisation and repeated practice of procedures to acquire mastery of mathematical concepts and skills. In contrast, close adherence to IMPACT would be likely to lead to lessons that emphasised pupil-led discussion, group work, and frequent use of concrete and pictorial materials before the introduction of the abstract. IMPACT teachers were expected to encourage collaborative problem-solving and guided discovery.[1]

The observation schedule comprised a mixture of rating types. Some variables would be expected to be observed very regularly in lessons, meaning that frequency was relevant. Examples include teacher-led instruction, or pupils engaged *on task*. For such variables, the observers indicated what percentage of each of a series of five-minute lesson segments was spent on each activity. The percentages of time within each segment were summed to produce an overall estimate of the percentage of observed class time for which each activity was observed.

The main activities rated in this way were:

- teacher providing instruction or information to pupils.
- pupils answering teacher questions.
- pupils asking questions or discussing maths (with others, or with teacher).
- pupils working as individuals (with or without help from teacher).
- pairs/small groups working on maths.
- time spent *on task*.

The percentages of time spent on the above activities do not sum to 100% as many overlap. For example, a teacher may be providing instruction while pupils are focused on task.

Another type of rating required the observers to draw on their professional experience as teachers to provide a more *global quality* rating. Examples of this type of rating include the extent to which differentiated teaching practices were evident during the lesson, and to which pupils seemed engaged. Finally, some ratings related to a global estimate of the amount of time spent on an activity. For example, observers provided an overall estimate of the percentage of class time pupils spent working from their workbooks/textbooks.

MQI ratings

All recorded lessons were observed by two subject matter experts (SMEs) and rated on a modified version of the MQI. In total, 60 sets of ratings were generated (30 lessons, each rated by two people). Lessons were rated as low, medium, or high on four dimensions:

- richness of the mathematics.
- working with pupils and mathematics.
- errors and imprecision.
- pupil participation in meaning-making and reasoning.

The dimension ratings were themselves based on a number of more specific indicators. For example, five items fed into the overall richness of the mathematics rating, while errors and imprecision and pupil participation were each based on three specific indicators, and working with pupils was based on two specific indicators.

[1] Of course, these expectations are based on the *relative*, not absolute, principles of JUMP and IMPACT. As outlined in Chapter 1, JUMP also promotes some guided discovery, although its definition of the term is different to that used generally. Equally, IMPACT promotes some scaffolding of discovery activities.

In most cases, a rating of high was a positive rating (e.g., teachers might reasonably be pleased if the *richness of the mathematics* in their lesson was rated as high). However, for errors and imprecision, a low rating was indicative of good quality teaching, as it meant that the teacher displayed few mathematical errors or instances of imprecise language. In addition to the four main dimensions, each SME rated lessons on two global dimensions (Mathematical Quality of Instruction [MQI] and Mathematical Knowledge for Teaching [MKT]).

SMEs reviewed and rated each lesson separately. Once ratings were completed, the two sets of ratings were compared. Given the quite limited nature of a three-point scale, it was unsurprising that SMEs varied on some ratings. Nonetheless, most ratings were identical or within a point of each other (e.g., one SME assigned a medium rating while another assigned a high). For the few occasions where there was a marked difference (low versus high), the SMEs and project researchers jointly reviewed the ratings and the lesson, coming to an agreed rating.

Lesson Report

SMEs watched each recorded lesson on a number of occasions. On the first occasion, they took notes and wrote a short lesson report. These reports described the lesson's structure, main activities, the time spent on each activity, and the details of the types of materials used. As it was expected that (if teachers adhered to programme principles) JUMP teachers would make less use of manipulatives and real life materials in a discovery context than would the IMPACT teachers, the lesson report notes these elements. As well as describing the lesson, the reports were drawn on by the SMEs when later rating specific aspects of the lesson.

Curriculum areas covered

The strands and strand units covered in each observed lesson (live or recorded) were noted.

Strand

In the first set of observations, Number was the strand most commonly covered, featuring in seven JUMP and 11 IMPACT lessons (Table 4.1). The next most common strand was Measures (four JUMP lessons and one IMPACT lesson), followed by Algebra, and Shape and Space. No lessons in either group dealt with Data. In contrast, by the second set of observations, almost all IMPACT lessons (10 of the 13) covered a Shape and Space topic, as did four JUMP lessons. However, other JUMP lessons covered topics in the Number, Measures and Data strands.

The marked popularity of Number, then Shape and Space topics among IMPACT teachers was probably related to the content of IMPACT manuals. The first IMPACT manual received at the start of the academic year dealt only with two Number strand units. The second manual, received closer to Christmas, dealt only with Shape and Space. While IMPACT principles are intended to be generalisable across the whole PSMC, it is likely that teachers felt more comfortable being observed while teaching content covered by one of the manuals.

Table 4.1: Number of JUMP and IMPACT classes in which various PSMC strands were taught during observed lessons

	JUMP (N=13)		IMPACT (N=13)	
	Time 1	Time 2	Time 1	Time 2
Number	7	3	11	1
Algebra	2	0	0	0
Shape and Space	0	4	1	10
Measures	4	3	1	2
Data	0	3	0	0

Strand unit

In the first observations, all Number strand units in JUMP classes related to either Place Value or Operations, the two Algebra lessons covered Number Patterns and Sequences, while the Measures strand units were Length (three lessons) and Time. In contrast, IMPACT mainly covered the Number strand units of Fractions (six lessons) and Decimals (four), with one lesson devoted to Operations. The only other strand units covered in IMPACT classes in the first set of observations were 3-D shapes (Shape and Space) and Length (Measures).

In the second set of observations, JUMP classes covered a wide variety of strand units. Three lessons covered Chance, two covered Fractions, with the following strand units covered in one of the observed lessons: Place Value, 2-D Shapes, 3-D Shapes, Symmetry, Capacity, Time, and Money. One lesson focused on coordinate points, which feature in JUMP but not the PSMC for Third class. IMPACT lessons were concentrated on the Shape and Space strand units of 2-D Shapes, 3-D Shapes, and Symmetry (10 lessons in total), with the three remaining lessons covering Fractions, Area, and Time.

Adherence to programme methods and principles

This section examines the extent to which lessons adhered to programme methods and principles, under seven broad headings:

- extent of teacher-led instruction (expected to be higher in JUMP classes).
- extent of pupil-led discussion (expected to be higher in IMPACT classes).
- pupil solo work (expected to be higher in JUMP classes) and group work (expected to be higher in IMPACT classes).
- types of materials used (workbooks expected to be more widely used in JUMP, and other materials expected to be more widely used in IMPACT).
- evidence of differentiated teaching practices (methods expected to differ by programme).
- learning styles (expected to differ by programme).
- use of assessment (expected to differ by programme).

The data presented are from the observation schedule, and relate to all 52 lessons observed.

Teacher-led instruction

Teacher-led instruction was operationally defined for observers as the teacher leading classroom instruction at whole-class level (including teacher-led question and answer sessions and work on the board, but excluding instruction of individual pupils or small groups while all pupils are carrying out solo or small-group work). For each of a series of five-minute time periods, observers rated the percentage of time spent in teacher-led instruction, with the percentages summed to gauge the percentage of total lesson time led by teachers.

For both sets of observations, a majority of class time in both programmes was taken up with teacher-led instruction (Table 4.2). However, the gap between programmes was smaller than might have been predicted during the first observations (averaging 67% of time in JUMP versus 58% in IMPACT lessons) and negligible during the second set of observations (57% in JUMP and 56% in IMPACT classes).

Question and answer sessions formed a subset of teacher-led instruction, and were defined as pupils answering questions *at whole-class level* (including listening to the teacher ask questions, waiting to answer questions [e.g., waving their hands] and actually answering questions). During the first observations, the two groups spent very similar percentages of class time on question and answer sessions (28% and 29% of JUMP and IMPACT lessons, respectively). By the second observations, the amount of time spent in teacher-led question and answer sessions had increased slightly for IMPACT lessons (to 36%), and increased noticeably for JUMP classes (to 41% of lesson time).

Table 4.2: Mean percentages of time spent on teacher-led instruction, questions and answers

	JUMP (N=13)		IMPACT (N=13)	
	Time 1	Time 2	Time 1	Time 2
Teacher-led instruction	66.8	56.7	58.1	55.9
Teacher-led Q&A	27.9	41.1	29.4	36.5

Observers also gave a general estimate of the amount of time that pupils spent listening to the teacher talk to or question the class. The estimates broadly mirrored the summed percentages, showing that in both groups (but especially in JUMP), most pupils spent considerable time listening to the teacher (Table 4.3).

Table 4.3: Estimates of time spent by pupils listening to the teacher talk to or question the class

	JUMP (N=13)		IMPACT (N=13)	
	Time 1	Time 2	Time 1	Time 2
Almost the entire lesson	1	2	0	0
Most of the time	7	5	6	5
About half the time	4	4	3	4
Some of the time	1	2	3	4
Hardly at all	0	0	1	0

Pupil-led discussion

Observers recorded the amount of time spent by pupils questioning or discussing mathematics, and listening to other pupils talk. They also rated the extent to which the classroom *climate* encouraged pupils to generate mathematical ideas and questions. As with teacher-led time, a series of five-minute ratings were summed to gauge the percentage of total lesson time spent in pupil-led discussion of mathematics. There were noticeable differences by programme, particularly during the first set of observations when only 7.0% of time in JUMP classes, but 31.4% of time in IMPACT classes, involved pupils questioning or discussing mathematics. By the second observations, this increased to 9.3% in JUMP, dropping to 25.1% in IMPACT classes.

Observers also gave a broad rating of the amount of time that pupils spent listening to other pupils talk. Although differences between programmes were less marked than for pupils discussing maths, the direction of difference was the same – i.e., pupils in JUMP classes were considered to have spent less time listening to other pupils talk than pupils in IMPACT classes (Table 4.4). For example, during the first set of observations, pupils in 10 JUMP and nine IMPACT classes spent either *some* or *hardly any* time listening to other pupils. By the second observations, this rose to 11 JUMP classes, while in IMPACT, it dropped to five classes.

Table 4.4: Estimates of time spent by pupils listening to other pupils talk

	JUMP (N=13)		IMPACT (N=13)	
	Time 1	Time 2	Time 1	Time 2
Almost the entire lesson	0	0	0	0
Most of the time	1	1	1	3
About half the time	2	1	3	5
Some of the time	4	9	8	5
Hardly at all	6	2	1	0

During the first set of observations only five JUMP (nine IMPACT) lessons included at least five minutes of group or class discussion of a mathematical task or question, but this increased to eight JUMP lessons (and 10 in IMPACT) by the second observations. The slight increase in the amount of pupil discussion in JUMP classes is reflected in a slight change in how the observers rated classroom climate (Table 4.5). Initially, no JUMP class was rated as *definitely* encouraging of pupil ideas and questions, while four were rated as *probably* encouraging. By the second observations, three classes were *definitely* encouraging, and a further three were rated as *probably* encouraging. That aside, IMPACT classes were generally rated as more encouraging than JUMP classes of pupil ideas.

Table 4.5: Numbers of observed classes in which classroom climate was rated as encouraging pupil ideas and questions, to various degrees

	JUMP (N=13)		IMPACT (N=13)	
	Time 1	Time 2	Time 1	Time 2
Definitely	0	3	3	4
Probably	4	3	2	4
Not sure	2	1	2	3
Not really	7	5	6	2
Not at all	0	1	0	0

Solo work and group work

Solo work was defined as pupils carrying out individual "seatwork" (e.g., working on workbooks or worksheets), including time spent by the teacher helping individual children. **Group work** was defined as time spent by pupils working in small groups or in pairs. There were large differences on these measures between the two programmes. On average, pupils in JUMP classes spent a relatively small amount of time in group work (14% and 11% for observations one and two, respectively), while IMPACT pupils spent close to half their lessons engaged in group work (42% and 43% for observations one and two, respectively) (Table 4.6). In a related vein, pupils in JUMP classes spent more time engaged in solo work – approximately 30% of time compared to 13% in IMPACT classes.

Table 4.6: Mean percentages of time spent on group and solo work

	JUMP (N=13)		IMPACT (N=13)	
	Time 1	Time 2	Time 1	Time 2
Group work	13.9	11.0	41.5	43.2
Solo work	29.4	30.9	13.9	13.2

In addition to differing in the mean percentage of time spent in group work, JUMP classes were less likely to include *any* significant amount of group activity (Table 4.7). Across the 26 JUMP lessons observed, only nine included at least five minutes of group or pair work (compared to 21 IMPACT lessons). In contrast, 21 of the 26 JUMP lessons observed included at least five minutes of independent individual work, as did 18 IMPACT lessons.

Table 4.7: Number of lessons in which independent individual work or pair/group work was observed for at least five minutes

	JUMP (N=13)		IMPACT (N=13)	
	Time 1	Time 2	Time 1	Time 2
Independent individual work	11	10	10	8
Independent pair or group work	4	5	11	10

Materials used

Observers were asked to indicate in which of a list of three types of activities (using textbooks/worksheets to answer questions, working with manipulatives, playing maths games) pupils were engaged for at least five minutes. On both occasions, the types of materials most widely used in JUMP were pupil questions in textbooks or workbooks, with less than half using manipulatives or maths games (Table 4.8). In contrast, manipulatives were used in most IMPACT lessons, with about half of classes also using maths games and questions in textbooks or workbooks.

Table 4.8: Number of classes that used particular materials/engaged in particular activities for at least five minutes

	JUMP (N=13)		IMPACT (N=13)	
	Time 1	Time 2	Time 1	Time 2
Textbooks/worksheets	9	10	6	6
Manipulatives	4	5	11	12
Maths games	5	4	6	7

Drawing on broader information from the lesson reports, a variety of manipulatives were used by classes in both programmes (e.g., “real life” materials such as dice, counters, rulers/metre sticks, and chocolate boxes, and structured materials such as cubes, number cards, geoboards, and Dienes blocks).

Observers’ estimates suggest that in addition to textbooks/worksheets being used in a larger number of JUMP classes, worksheets (or workbooks, in the JUMP programme) were used for longer periods of time. For example, during the first observations, textbooks/worksheets were used for 41-60% of lesson time in five JUMP but only two IMPACT classes (Table 4.9). During the second set of observations, eight JUMP classes used workbooks for at least 20% of class time, compared to only two IMPACT schools.[2]

[2] The amount of time in IMPACT schools may be slightly underestimated, as (unlike Table 4.8) it does not include time pupils spent copying questions from the textbook/the board and answering in their copybooks.

Table 4.9: Number of observed classes that spent within various percentage ranges of class time using worksheets/workbooks

	JUMP (N=13)		IMPACT (N=13)	
	Time 1	Time 2	Time 1	Time 2
61%+	0	0	0	0
41-60%	5	4	2	1
21-40%	1	4	1	1
1-20%	3	0	3	2
None	4	5	7	9

Differentiated teaching

As outlined in Chapter 1, bonus questions are proposed by the JUMP programme as a means of differentiating, since they should allow higher-achieving pupils to work independently while giving the teacher time to assist lower-achieving pupils by breaking down concepts and skills. The collaborative problem-solving methods advocated by IMPACT (including questioning and re-voicing techniques, and use of varied models) should also promote differentiated teaching, since they aim to allow pupils at various levels and with various learning styles to participate in developing solution methods. Therefore, observers provided an overall rating of the extent to which differentiated teaching practices were evident in the observed lesson, and, more specifically, if bonus questions and collaborative problem-solving were used in the lesson.

Clear use of differentiated teaching practices was observed in only a minority of lessons, most of which were IMPACT classes (Table 4.10). During the first set of observations, six JUMP lessons were rated as *not really* displaying differentiated practice, with an additional class where differentiated teaching was *not at all* present. By the second set of observations, five JUMP lessons were described as *not at all* showing differentiated practice, with a further four *not really* showing evidence of differentiation.

Table 4.10: Number of observed classes in which differentiated teaching was deemed present, to various degrees

	JUMP (N=13)		IMPACT (N=13)	
	Time 1	Time 2	Time 1	Time 2
Definitely	0	1	2	5
Probably	3	2	2	0
Not sure	3	1	4	1
Not really	6	4	4	3
Not at all	1	5	1	4

Despite the relatively poor observer ratings for differentiation, bonus questions were used in nine JUMP classes during the first set of observations, falling to only five during the second set (Table 4.11). A teacher in one IMPACT class used a method similar to bonus questions during both observations. Thus, bonus questions were used less than might be anticipated in JUMP lessons, and, where used, were not always used in a manner that observers believed represented differentiated teaching practice.

Table 4.11: Number of observed classes in which bonus questions were given to pupils

	JUMP classes (N=13)		IMPACT (N=13)	
	Time 1	Time 2	Time 1	Time 2
Bonus questions used	9	5	1	1

Collaborative problem-solving was not a feature of most JUMP classes during either set of observations, with observers rating it as *not really* or *not at all* manifested in a majority of classes (Table 4.12). However, observers also reported that it was *probably* or *definitely* a feature of only half of IMPACT classes.

Table 4.12: Number of observed classes in which collaborative problem-solving was deemed present, to various degrees

	JUMP (N=13)		IMPACT (N=13)	
	Time 1	Time 2	Time 1	Time 2
Definitely	0	1	4	4
Probably	2	1	3	2
Not sure	1	2	0	1
Not really	5	4	4	0
Not at all	5	5	2	6

Learning styles

As noted in Chapter 1, a central tenet of the JUMP philosophy is that repetition and practice are key to learning. Therefore, observers were asked to report the extent to which these were present in the observed lessons. In the first set of observations**,** memorisation and repetition of procedures were either *definitely* or *probably* present in 11 JUMP classes, but also in four IMPACT classes (Table 4.13). However, during the second observations, JUMP classes were reasonably evenly split between those who showed some or no evidence of memorisation and repeat procedures.

Table 4.13: Number of observed classes in which memorisation and repeat procedures were deemed present, to various degrees

	JUMP (N=13)		IMPACT (N=13)	
	Time 1	Time 2	Time 1	Time 2
Definitely	7	5	4	2
Probably	4	2	0	2
Not sure	0	0	1	1
Not really	2	3	6	2
Not at all	0	3	1	6

Observers ranked the broad approaches of “simple direct instruction (explanation)”, “instruction by a series of related questions” and “guided discovery (activities)”, in order of their frequency of use in each lesson (Table 4.14). During both sets of observations, JUMP classes tended to use instruction by a series of questions or simple direct instruction, with guided discovery the least used strategy in almost all classes. In IMPACT classes, instruction by a series of questions was also the most common strategy, but in contrast to JUMP, guided discovery was the second most common strategy.

Table 4.14: Number of observed classes in which simple direct instruction, instruction by a series of questions, and guided discovery were rated as used with highest, medium, or lowest frequency

		JUMP (N=13)		IMPACT (N=13)	
		Time 1	Time 2	Time 1	Time 2
Simple direct instruction	Highest	6	6	4	1
	Medium	5	6	3	4
	Lowest	2	1	6	8
Series of questions	Highest	7	7	5	7
	Medium	6	6	6	5
	Lowest	0	0	2	1
Guided discovery	Highest	0	0	4	5
	Medium	2	1	4	4
	Lowest	11	12	5	4

Assessment

The observation schedule did not include assessment-related activities, due to the difficulty in pre-defining assessment practices in a manner that could be uniformly understood and rated. Instead, the SMEs' reports on the recorded classes provided data on how assessment was used in those 15 classes (30 lessons).

In both sets of observations, all recorded classes featured some form of continuous assessment. However, this was often limited to a teacher's monitoring of pupils' solo work. At other times, assessment was combined with oral review of material. In both these situations, it was usually the case that only *some* work of *some* pupils could be assessed. (That said, in two JUMP and two IMPACT classes within the recorded group, small class size or team teaching meant that *all* pupils could be informally assessed throughout the lessons).

In the first observations, more distinct forms of summative assessment were observed in three of the seven JUMP classes. In one class, pupils were given two worksheets as "quizzes". Each took approximately four minutes to complete, and all quizzes were collected by the teacher for correction. In the other two classes, approximately four minutes were spent checking answers at whole-class level, using the board. Two of the eight recorded IMPACT classes also devoted time to distinct forms of assessment. In one case, approximately 10 minutes were spent eliciting and discussing each pupil's solution to a problem (formative assessment). In the other case, two segments of approximately three minutes each were spent checking answers at whole-class level, using the board. Thus, programme-related patterns of assessment did not emerge strongly in the JUMP and IMPACT groups during the first observations.

In the second set of observations, there was even less evidence of distinct forms of assessment, although most JUMP and some IMPACT lessons included some oral review questions at whole-class level. One JUMP lesson also included a written review test, corrected verbally by the teacher at the start of the lesson, while another included time spent correcting workbook answers on the board. In general, however, assessment methods did not appear to differ notably by programme.

Global ratings of adherence to programme

For each lesson, observers rated on a scale of 1-10 how closely it adhered to the assigned programme. During the first set of observations, observers were asked to rate JUMP lessons for adherence to the selected lesson plan(s). For the second set of observations, observers provided two ratings of JUMP lessons – the degree of fidelity to the lesson plan and the degree of fidelity to JUMP principles. Ratings were split for the second observations because it was assumed that teachers might by that late stage in the year have become more comfortable with JUMP principles and proficient in JUMP methodologies. As such, they might be more likely to diverge from a set JUMP lesson plan, while still adhering to its general principles. Since the IMPACT programme did not contain lesson plans per se, global ratings for IMPACT were based on adherence to principles during both observations.

Most teachers in each group showed some levels of adherence to the relevant lesson plan or programme principles, but few showed very high levels. For JUMP, the average adherence rating to the lesson plan was 7.0 (out of 10) during the first set of observations, rising marginally to 7.2 for the second set (Table 4.15). With an average rating of 7.5, adherence to JUMP *principles* was slightly higher, although one lesson only received a rating of four out of 10. With IMPACT, average ratings for adherence to programme principles dropped between the first and second observations (from 7.2 to only 6.7). However, this can be attributed to one teacher, whose adherence ratings dropped from an initial rating of eight to two during the second observed lesson.

Table 4.15: Observers' overall ratings, on a scale of one (low) to 10 (high), of adherence to programmes (lesson plan and/or principles) in observed JUMP and IMPACT classes

	JUMP (N=13)			IMPACT (N=13)	
	Time 1	Time 2		Time 1	Time 2
	Lesson plan	Lesson plan	Principles	Principles	Principles
Mean rating	7.0	7.5	7.2	7.2	6.7
Range	5-9	5-9	4-9	5-9	2-9

Quality of instruction

To rate the *quality* of instruction in observed classes, the observation schedule (applied to all 52 observed lessons) included measures of pupil engagement and pupil understanding, while the subset of lessons which were recorded were also rated using a modified version of the MQI.

Pupil engagement

Three ratings of pupil engagement with the lesson were provided. First, observers indicated if they felt that *most or all* pupils were engaged with the lesson. Second, they estimated the proportion of pupils likely to have had a good understanding of the lesson by the end of class, and third, they rated the amount of time pupils spent "on task". Time on task is generally taken as meaning the amount of time in which pupils were actively engaged with their school work. In this instance it included pupils actively working on a task, paying attention to a teacher or a classmate, and the time taken to set up an activity *efficiently*.

Generally, pupil engagement with the lesson was high. Of the 26 JUMP lessons observed, pupils were rated as *definitely* or *probably* engaged with the lesson in all but one instance (Table 4.16). Engagement was also generally high in IMPACT classes, although in four lessons, observers were unsure of pupil engagement levels, and in one case pupils were rated as *not really* engaged.

Observers were slightly less positive when rating pupil understanding of lesson content. In the first observations, none of the 13 JUMP classes were rated as having all pupils possess a *good* understanding of the topic by the end of the lesson. However, in 10 classes, observers considered that over half the pupils had a good understanding. During the second observations, three observers felt that the entire class developed a good understanding of the lesson topic, but in four classes no more than about half of pupils were believed to have such an understanding. IMPACT classes showed a broadly similar pattern in that in most cases *over half* of pupils were believed to have a good understanding of the topic.

Table 4.16: Number of observed classes in which most or all pupils were considered engaged to various degrees, and in which various proportions of pupils had a good understanding of lesson content

		JUMP (N=13)		IMPACT (N=13)	
		Time 1	Time 2	Time 1	Time 2
Most/all pupils engaged with lesson	Definitely	4	5	5	4
	Probably	9	7	5	7
	Not sure	0	1	2	2
	Not really	0	0	1	0
	Not at all	0	0	0	0
Pupils with good understanding of content	All of them	0	3	3	1
	Over half	10	6	5	8
	About half	2	3	4	3
	Less than half	1	1	0	0
	None	0	0	0	0

Time on task was assessed in two ways, both of which suggested that for the vast majority of lesson time, pupils were generally on task. First, based on summed estimates from a series of five-minute lesson segments, JUMP pupils were on task 96.3% of the time during the first series of observations, and 94.3% of the time for the second set. For IMPACT, the equivalent data were 90.0% and 91.9%, respectively. Second, at the end of each lesson, observers provided a general estimate of how much time they felt was on task. Five options were presented, ranging from *less than 25%* to *over 90%* of time, but only the two highest were used by observers (Table 4.17). During both observations, a sizeable majority of JUMP classes were rated as spending over 90% of time on task. Ratings were also quite positive for IMPACT, although during the first set of observations six of the 13 lessons were rated as 76-90% on task.

Table 4.17: Number of classes in which percentage of class time spent on task was estimated to fall within the 91-100% or 76-90% ranges

	JUMP (N=13)		IMPACT (N=13)	
	Time 1	Time 2	Time 1	Time 2
91%+	9	10	7	9
76-90%	4	3	6	4

Mathematical Quality of Instruction (MQI)

The MQI (Learning Mathematics for Teaching Project, 2011) was used to rate all recorded observations on four broad dimensions (richness of mathematics, working with pupils and mathematics, teacher errors and imprecision, and pupil participation in meaning-making). Each of 15 class groups was recorded on two occasions, and rated by two SMEs, making a total of 60 sets of MQI ratings for the 30 individual lessons recorded.

The remaining tables in this chapter show SME ratings by programme and by observation. As will be seen, almost every rating improved between the first and second observations. This may be attributable to participation in the programmes improving quality of instruction, or to teachers generally reflecting more on their teaching practice as a result of participation in an evaluation. However, many teachers informally indicated that they were more comfortable in front of the camera on the second occasion, and less constrained by nervousness. It is therefore also possible that the data in subsequent tables may reflect this rather than any programme or evaluation effects.

Richness of the mathematics

Instruction was defined as featuring rich mathematics if it was focused on the meaning of mathematical facts and procedures, and/or deeply engaged with mathematical practices and language. The richness of the mathematics in a lesson was rated as low (1), medium (2), or high (3) on five sub-dimensions, which were then used to inform the overall rating for richness of mathematics. The five sub-dimensions were:

- linking and connection (e.g., of ideas, procedures, representations).
- teacher explanations (e.g., of why a procedure works, why a solution makes sense).
- multiple procedures or solution methods (for a single problem or a problem type).
- developing mathematical generalisations (e.g., examining cases and noting a pattern).
- mathematical language (fluent and explicit use of mathematical terms).

Comparing JUMP and IMPACT ratings, the sub-dimension on which they differed most notably was teacher explanations (Table 4.18). JUMP classes initially received a mean rating of 2.5 (i.e., medium to high quality) for teacher explanations, which rose to 2.8 for the second observations. Ratings for teacher explanations in IMPACT lessons also increased, but from an average of 1.8 (just below medium quality) to 2.3 (a little above medium).

Table 4.18: Observers' mean ratings of recorded lessons on the Richness of the Mathematics dimension, on a scale of 1 (Low) to 3 (High)

	JUMP (N=14 ratings, 7 classes)		IMPACT (N=16 ratings, 8 classes)	
	Time 1	Time 2	Time 1	Time 2
Teacher Explanations	2.5	2.8	1.8	2.3
Mathematical Language	1.8	2.3	1.6	2.4
Linking and Connection	1.7	2.0	1.6	2.0
Multiple Procedures or Solution Methods	1.3	1.5	1.5	1.4
Developing Mathematical Generalisations	1.3	1.7	1.3	1.5
Overall Richness of the Mathematics	1.6	2.1	1.4	2.0

Ratings for teacher use of mathematical language in JUMP lessons increased from 1.8 to 2.3, but from 1.6 to 2.4 in IMPACT. Across both programmes, ratings were lowest for the use of multiple procedures and solution methods and for developing mathematical generalisations. The overall richness of the mathematics was estimated at a mean of 1.6 in JUMP and 1.4 in IMPACT (i.e., low to medium for both groups) for the first observations, but rose to 2.1 and 2.0, respectively, for the second observations.

Working with pupils and mathematics

In assessing how appropriately teachers responded to pupils' mathematical errors and productions, the two sub-dimensions rated were:

- remediation of pupil errors and difficulties (at procedural and conceptual levels).
- responding to pupil mathematical productions in instruction (e.g., identifying the mathematical relevance of pupil questions, using pupil ideas to build instruction).

As with the previous dimension, SME ratings suggested an improvement in quality between the first and second observations. The improvement was slightly more pronounced in IMPACT ratings, especially in how teachers responded to pupil mathematical productions – rising from 1.6 (low to medium) to a mean rating of 2.3 (medium to high) by the second observations (Table 4.19). On the overall scale for working with pupils and mathematics, the mean rating for JUMP classes was 2.1 (medium quality) on both occasions. For IMPACT, it was 1.7 (a little below medium quality), which rose to 2.0 (medium) for the second observations.

Table 4.19: Observers' mean ratings of recorded lessons on the Working with Pupils and Mathematics dimension, on a scale of 1 (Low) to 3 (High)

	JUMP (N=14 ratings, 7 classes)		IMPACT (N=16 ratings, 8 classes)	
	Time 1	Time 2	Time 1	Time 2
Remediation of Pupil Errors and Difficulties	1.8	2.0	1.4	1.8
Responding to Pupil Mathematical Productions	2.1	2.4	1.6	2.3
Overall Working with Pupils and Mathematics	2.1	2.1	1.7	2.0

Errors and imprecision

For the dimension of errors and imprecision, a low rating (1) was a positive rating (i.e., low level of teacher error), while a high rating of 3 was a negative rating. There were three sub-dimensions to the error and imprecision rating:

- major mathematical errors (e.g., solving problems incorrectly, omitting a key condition in a definition).
- imprecision in language or notation (e.g., errors in mathematical symbols or language).
- lack of clarity in presentation of mathematical content (e.g., mathematical point is muddled).

In the first set of observations, all teachers were rated as displaying low levels of major mathematical errors (i.e., a positive rating) (Table 4.20). Ratings were also almost uniformly positive for imprecision in mathematical language or notation, but slightly less positive for IMPACT teachers where clarity of presentation was concerned. Six (of the eight) IMPACT teachers were rated by at least one SME as having medium levels of clarity in presentation. Overall, though, the SMEs' ratings for errors and imprecision indicate low levels of errors evident during the first set of observations.

In the second set of observations, JUMP teacher ratings indicate a slight increase in errors and imprecision. For example, two JUMP teachers were rated by both SMEs as displaying a medium level of mathematical errors during the lesson (compared to no teachers during the first observations). The mean rating for lack of clarity also increased for JUMP lessons, while decreasing slightly for IMPACT. The overall errors and imprecision rating for JUMP lessons for the second observations was 1.2, compared to 1.1 for IMPACT lessons.

Table 4.20: Observers' mean ratings of recorded lessons on the Errors and Imprecision dimension, on a scale of 1 (Low level of error) to 3 (High level of error)

	JUMP (N=14 ratings, 7 classes)		IMPACT (N=16 ratings, 8 classes)	
	Time 1	Time 2	Time 1	Time 2
Major Mathematical Errors	1.0	1.3	1.0	1.1
Imprecision in Mathematical Language/Notation	1.1	1.3	1.1	1.2
Lack of Clarity in Presentation of Content	1.1	1.4	1.4	1.2
Overall Errors and Imprecision	1.1	1.2	1.1	1.1

Pupil participation in meaning-making and reasoning

Pupils were considered to participate in meaning-making and reasoning when they provided explanations, generated questions or arguments, and demonstrated engagement at a high cognitive level. Three sub-dimensions were rated, to contribute to the overall dimension:

- pupils provide explanations (may be pupil-initiated or teacher-initiated).
- pupil mathematical questioning and reasoning (e.g., pupils ask questions requiring explanations, make conjectures, or reason out conclusions).
- enacted task cognitive activation (i.e., whether pupils engage with tasks using a low, mixed or high level of thinking skills).

In the first observations, JUMP lessons were rated slightly lower than IMPACT lessons on pupils providing explanations (1.5 versus 1.8) (Table 4.21). By the second observations, this difference was reversed, with JUMP lessons averaging a rating of 1.9 (close to medium) for pupils providing explanations, compared to a rating of 1.7 for IMPACT lessons. Both groups received initial mean ratings of 1.4 for pupil mathematical questioning and reasoning (low to medium levels), and ratings for both rose slightly for the second observations.

Table 4.21: Observers' mean ratings of recorded lessons on the Pupil Participation in Meaning-Making and Reasoning dimension, on a scale of 1 (Low) to 3 (High)

	JUMP (N=14 ratings, 7 classes)		IMPACT (N=16 ratings, 8 classes)	
	Time 1	Time 2	Time 1	Time 2
Pupils Provide Explanations	1.5	1.9	1.8	1.7
Pupil Mathematical Questioning and Reasoning	1.4	1.6	1.4	1.7
Enacted Task Cognitive Activation	1.9	2.0	1.6	2.1
Overall Pupil Participation in Meaning-Making/Reasoning	1.6	1.9	1.7	2.2

JUMP lessons were initially rated slightly higher than IMPACT lessons for enacted task cognitive activation (1.9 versus 1.6). For the second observation, JUMP lesson ratings increased marginally, whereas IMPACT lessons rose by half a point (to just above medium). Initially, the ratings for overall pupil participation in meaning-making and reasoning were similar (1.6 for JUMP and 1.7 for IMPACT). However, in the second observations, IMPACT lessons averaged 2.2 (a little above a medium rating) whereas JUMP averaged 1.9 (or just below medium).

Global ratings of MQI and MKT

For each lesson viewed, each SME gave an overall estimate of mathematical quality of instruction as either low, medium or high. They also took a "lesson-based guess" at whether a teacher's mathematical knowledge for teaching (MKT) was low, medium or high. For both programmes, the overall average MQI rating for teachers increased from the first to the second observation. MKT estimates remained static for JUMP, but the SMEs rated IMPACT teachers' mathematical knowledge for teaching higher in the second observation than they did in the first (Table 4.22). For the second observations, quality of instruction was close to *medium* for both programmes. Teacher MKT was estimated as a little below medium (mean of 1.7) for JUMP teachers and medium (mean of 2.0) for IMPACT teachers.

Table 4.22: Observers' ratings of recorded lessons for Mathematical Quality of Instruction and estimated Mathematical Knowledge for Teaching scores, on a scale of 1 (Low) to 3 (High)

	JUMP (N=14 ratings, 7 classes)		IMPACT (N=16 ratings, 8 classes)	
	Time 1	Time 2	Time 1	Time 2
MQI	1.7	1.9	1.5	2.1
MKT	1.7	1.7	1.6	2.0

General comments on lessons

As noted earlier, a lesson report was completed by an SME for each recorded class. The lesson reports, in conjunction with informal conversations with the SMEs and those who conducted the live observations, provided supplementary qualitative data on how lessons were enacted in practice.

The SMEs noted that review-style lessons were common in both programmes and both sets of observations, but particularly in the first set. However, few lessons in either group ended with overall recaps of lesson content. A partial explanation may be that some teachers reported being nervous during the first recorded observations, but being less so during the second (also evident in higher MQI scores for the second observation). Teaching new material might be seen as more stressful than reviewing familiar material, suggesting that the emphasis on review was related more to nerves than to adherence to a review-heavy teaching approach.

The lesson reports also raised concerns that teachers sometimes adhered closely to the letter, but not the spirit, of their assigned programme. In JUMP, some lessons featured considerable repetition and practice, but lacked reference to the larger mathematical ideas behind the repeated steps. In a small number of cases, the teacher was so intent on the "step-by-step" approach that they rejected suggestions from pupils who had moved ahead or come up with an alternative (correct) solution method. One SME noted that the high level of decomposition of concepts and skills that *should* characterise JUMP lessons might be considered a form of differentiation to be used with weaker pupils. However, the fact that relatively little differentiation was evident was an indicator of low adherence to some key JUMP principles.

In IMPACT, some lessons drew heavily on concrete materials, but did not link the concrete practice with abstract mathematical concepts. In others, the social constructivist approach to analysing a problem and generating a shared solution seemed poorly understood. While the IMPACT programme promotes the idea that the teacher should not be the sole validator of knowledge and that all pupil responses should be valued, a few teachers' interpretation of this was to treat all pupil responses as equally *correct*. Several lessons featured pupils talking at length about problems, but the construction of a solution was not a strong feature of the discussion.

In both groups, almost half the lessons started and/or ended with a session of mental mathematics practice, often unrelated to the main lesson topic. In JUMP classes, this was typically achieved through fast-paced exercises and games using an interactive whiteboard. In IMPACT classes, it sometimes involved activities specifically recommended by the programme, e.g., the "counting choir" and "sound of a number" (counting can). The SMEs noted that this approach has been endorsed by many within the Inspectorate and that teachers may be trying to use mental mathematics as a means of improving number understanding and skill.

The lesson reports also showed that superficially similar teaching strategies could have quite different quality ratings. For instance, "station teaching" was used in several IMPACT classes in both sets of observations.[3] However, the SMEs viewed it as having less positive outcomes when used throughout an entire lesson than when used in between sessions of teacher-led introduction and recapping. They also felt that the amount of time allocated per station should (but did not always) take account of the task complexity and the amount of time needed for set up or explanation of the task. While station activities were usually drawn from the IMPACT manual, they were sometimes presented to pupils without context – a further instance of adhering to the letter of the programme, but not to its spirit.

Further, pupils in a few lessons seemed overly familiar with the station activities, suggesting that these observed lessons might have been repeated lessons. Indeed, in most cases it would have been impossible to run station-based lessons without an initial introduction, as pupils would not know what to do at each station. Some of those who conducted the live observations felt that at least some aspects of a small number of lessons had been rehearsed. Supporting evidence for this is found in some pupils' comments during interviews. For instance, when asked whether the lesson was typical, one pupil remarked that it was "*the same as most classes since two weeks ago*".

In a few other cases, observers inferred from conversation with teachers and pupils that the lessons they watched were atypical in their adherence to the assigned programmes. Some teachers commented that while they did not usually stick rigidly to their programme, they had made a particular effort for the observation. As will be outlined in Chapter 5, some pupils noted that the amount of games and activities used in observed lessons was unusual.

On balance, it may be that the observed lessons (particularly the first set) did not all represent typical mathematics lessons in the classes observed. A small number may have repeated certain elements from previous lessons, while others may have contained "*more props than normal*" (as one pupil succinctly put it). In a separate but related vein, the possibility that there could have been a similar approach to the end-of-year testing cannot be excluded. In two classes, observers expressed a concern that pupils might have been schooled on the DPMT.

[3] Although the IMPACT programme suggests activities for small group work, the manuals do not explicitly mention station teaching.

Summary

The observations were intended to provide measures of teachers' adherence to assigned programmes, and to examine the quality of instruction experienced by pupils in the two groups. Observer ratings indicated moderately strong adherence to assigned programmes. Most teachers in each group demonstrated *some* adherence to the relevant lesson plan or programme principles, but few demonstrated very close adherence, and a few appeared not to be basing their lesson on the assigned programme principles. Data from the observation schedule suggest that there was higher adherence to certain aspects of the programmes. In particular, the emphasis on teacher-led instruction and solo work was, as would be expected, more prominent in JUMP than IMPACT lessons, and group work and pupil discussion were more prominent in IMPACT than JUMP. However, prominence was relative to the other programme, and in most cases, not markedly high in a broader sense. For example, JUMP classes had more solo work than IMPACT classes, but not enough solo work to be considered outside the normal range of classroom practices.

Other aspects of programmes were less frequently observed than might be expected, particularly during the observations conducted later in the school year (e.g., JUMP bonus questions, use of memorisation and repeat procedures, use of workbooks). This, in conjunction with already noted concerns as to the representativeness of a few of the first observations, suggests that aspects of programmes might have been patchy in some classes. At a more qualitative level, the SMEs commented that in the recorded observations, there was relatively little evidence of the IMPACT programme's emphasis on a social constructivist approach or JUMP's intended reliance on frequent use of summative assessment.

Ratings of quality of instruction were similar across the two groups. Pupil engagement and *time on task* were high overall, while pupil understanding was slightly lower. Generally, quality ratings improved between the first and second set of recorded observations, perhaps due to teachers becoming more comfortable in front of a camera. JUMP and IMPACT lessons received similar ratings on most MQI dimensions. A notable exception was the quality of teacher explanations, which was higher on both occasions in JUMP lessons. Informal comments from SMEs and observers suggested that some observed lessons, particularly those in the first set, may not have always been an accurate reflection of a typical lesson.

Chapter 5
Pupils

This chapter examines characteristics of the pupils who took part in the two programmes, drawing on questionnaire and interview data. Pupil Questionnaires were administered to all pupils in September 2013 and at the end of May 2014. To ensure comparison within the same group of pupils, data are reported only for those who completed a questionnaire on both occasions. Interview data are drawn from interviews with subsets of pupils in December 2013 and May 2014. For the interviews, a total of 78 pupils (39 from each programme) were interviewed on each occasion. For each class, a group interview was conducted with three mixed-ability pupils. Pupils were selected by teachers, and in many cases, different pupils were selected in December and May. Thus, while the questionnaire data are derived from the same pupils on both occasions, the interview data relate to two different but overlapping sets of pupils.

Based on the 509 pupils who completed both Pupil Questionnaires, almost all pupils in each programme were either eight or nine years old, and there were slightly more boys than girls taking part, particularly in JUMP, where 58% of pupils were boys (Table 5.1).

Table 5.1: Summary gender and age information for pupils completing the Pupil Questionnaires

		% JUMP (N=271)	% IMPACT (N=238)
Gender	Boys	58.3	54.2
	Girls	41.7	45.8
Age	7	0.4	1.3
	8	56.7	48.9
	9	41.5	48.1
	10	1.5	1.7

Information on gender, but not on age, was collected from pupils who took part in the group interviews. Unlike the composition of the overall population of pupils, the gender split was slightly in favour of girls. During the first set of interviews, 40 of the 78 pupils spoken to (51.3%) were girls, as were 41 (52.6%) pupils who took part in the second set of interviews.

The rest of this chapter examines pupil attitudes and experiences under five main headings:

- Attitudes to school.
- Attitudes to mathematics.
- Confidence and anxiety in relation to mathematics.
- Experiences of classroom practice in relation to mathematics.
- Strategies used in mathematics lessons.

Attitudes to school

In both programmes, less than half of pupils indicated that they liked school, while approximately one third were unsure of their opinion (Table 5.2). In September 2013, a higher percentage of IMPACT than JUMP pupils indicated that they disliked school (27% versus 18%, respectively), a gap that remained largely unchanged in the May administration of the questionnaire.

Table 5.2: Percentages of pupils indicating if they liked school

	September 2013		May 2014	
	% JUMP (N=265)	% IMPACT (N=232)	% JUMP (N=268)	% IMPACT (N=231)
Like school	46.4	40.1	47.8	32.9
Not sure	35.1	33.2	32.8	39.8
Do not like school	18.5	26.7	19.4	27.3

For both programmes and on both occasions, boys were far more likely than girls to say they disliked school. For example, no more than 10% of girls in any programme on either occasion indicated that they did not like school. However, for boys in the IMPACT group, 41% surveyed in September indicated they did not like school, rising slightly to 43% in May.

Attitudes to mathematics

In the questionnaire, pupils were asked the extent to which they agreed with a number of statements relating to mathematics, using a four-point scale (*agree a lot, agree a little, disagree a little, disagree a lot*). Table 5.3 presents summary information on responses, showing combined *agree a lot* and *agree a little* responses. As can be seen, most pupils expressed positive attitudes to mathematics, irrespective of programme. Many attitudes changed little across the two time points. In particular, the percentages who believed that they were good at maths were almost identical on both occasions, while there was minimal change in the percentages who worried about being asked questions in class. There were, however, slight increases in the percentages indicating that they liked maths (by 6% in both JUMP and IMPACT), who learned interesting things in maths lessons (by 8% in JUMP and 3% in IMPACT), and who believed everyone could be good at maths (a 5% increase among JUMP pupils and a 7% increase among IMPACT pupils).

Table 5.3: Percentages of pupils who *agreed a lot/little* with various written statements about mathematics

	September 2013		May 2014	
	% JUMP (N=271)[1]	% IMPACT (N=238)	% JUMP (N=271)	% IMPACT (N=238)
I like maths	69.4	63.8	75.7	69.7
I wish I didn't have to study maths	45.9	48.0	43.3	44.7
I learn interesting things in maths	74.2	78.4	82.5	81.2
I am good at maths	81.0	76.2	81.2	76.7
I think everyone can be good at maths	80.1	67.7	85.4	74.9
I worry that I won't be able to answer questions in maths class	47.7	49.8	45.9	49.8

[1] To facilitate simplified presentation of data, *N*s shown are the maximum number in each programme who responded to a statement. Actual *N*s vary slightly by statement (ranging from one to 15 missing responses).

When interviewed, the first question pupils were asked was if they had a favourite subject, and if yes, was it mathematics. In the December interviews, slightly more than half of the pupils in each programme agreed that mathematics was their favourite subject (Table 5.4). However, in May, only 29% of JUMP and 26% of IMPACT pupils interviewed agreed this was true – a sizeable drop. Nonetheless, JUMP pupils interviewed were more positive about liking mathematics than the questionnaire responses suggested. During both the December and May interviews, 87% of JUMP pupils said that they liked maths, and no pupils said they did not like the subject. In contrast, between the December and May interviews, the percentage of IMPACT pupils liking mathematics dropped from 79% to 66%, with 29% expressing ambivalence during the May interviews.

Table 5.4: Percentages of pupils who agreed, disagreed or were unsure if they liked maths or it was their favourite subject

			Yes	Unsure	No
Is maths your favourite subject?	Dec 2013	JUMP (N=37)	56.8		43.2
		IMPACT (N=31)	51.3		48.7
	May 2014	JUMP (N=39)	29.0		71.0
		IMPACT (N=38)	26.3		73.7
Do you like maths?	Dec 2013	JUMP (N=39)	86.8	13.2	0.0
		IMPACT (N=39)	79.5	12.8	7.7
	May 2014	JUMP (N=37)	86.5	13.5	0.0
		IMPACT (N=38)	65.8	28.9	5.3

Pupils were asked if they preferred mathematics when they were in Second class or if they preferred it at the time of interview. During the December interviews, a large majority (89% in JUMP AND 82% in IMPACT) preferred their current experience of mathematics. However, when interviewed towards the end of the school year, only 54% of IMPACT pupils preferred their current mathematics lessons, with 72% indicating that they felt mathematics lessons were more fun in Second class. Most JUMP pupils remained positive, with 82% preferring their current mathematics lessons to those in Second class.

Gender differences in attitudes to mathematics

In both start- and end-of-year Pupil Questionnaires, girls in both programmes were more likely than boys to report liking maths (Table 5.5). The percentages that liked maths increased for both genders from September to May (almost identically for girls and boys in JUMP, and just 2% more for boys than girls in IMPACT). However, there was a comparatively large increase (10%) in the percentage of JUMP boys reporting they learned interesting things in maths (compared to 6% for JUMP girls, and 3% for girls and boys in IMPACT). The slight overall drop in percentages of pupils wishing they didn't study maths is mainly attributable to JUMP boys (a decrease of 4%) and IMPACT girls (a decrease of 9%).

In September, percentages of pupils agreeing that they were good at maths were similar for girls and boys in JUMP, and 7% higher for boys than girls in IMPACT.[2] In May, the percentages agreeing they were good at maths increased for JUMP boys and IMPACT girls (4% in each case), while they decreased for JUMP girls (5%) and IMPACT boys (2%). Percentages agreeing that everyone could be good at maths increased across all groups, but the largest

[2] However, in both programmes and on both occasions, higher percentages of boys than girls *agreed a lot* (as opposed to *a little*) with this statement.

increases were for JUMP boys (7%) and IMPACT girls (14%). There was little change for any group in percentages worrying about being asked questions in class.

In sum, the Pupil Questionnaires suggest that positivity towards mathematics and confidence in mathematical ability increased more substantially for boys than girls in JUMP, and for girls than boys in IMPACT. However, most of these percentage differences remain small.

Table 5.5: Percentages of pupils who *agreed a lot/little* with various written statements about mathematics, by gender

	September 2013				May 2014			
	JUMP		IMPACT		JUMP		IMPACT	
	Girls (N=113)	Boys (N=158)	Girls (N=109)	Boys (N=129)	Girls (N=113)	Boys (N=158)	Girls (N=109)	Boys (N=129)
I like maths	72.0	67.6	66.6	61.5	78.3	73.8	71.3	68.3
I wish I didn't have to study maths	44.1	47.1	51.4	45.2	44.3	42.6	42.2	46.9
I learn interesting things in maths	74.1	74.2	81.0	76.2	80.2	84.1	83.5	79.2
I am good at maths	81.6	80.5	72.1	79.5	77.0	84.2	76.0	77.3
I think everyone can be good at maths	86.9	75.2	65.3	69.7	90.0	82.2	79.5	70.7
I worry that I won't be able to answer questions in maths class	55.1	42.6	55.7	44.8	54.9	39.4	56.9	43.8

Favourite and least favourite aspects of mathematics

Interviewers asked pupils to name "one best thing and one worst thing about learning maths". Table 5.6 shows the most popular pupil responses supplied. A large number of answers focused on specific strand units or elements thereof (e.g., "*I don't like doing division*" or "*I love doing multiplication sums*"). Across the two set of interviews, pupils in the JUMP programme were more likely to cite mathematics as being fun or to say they liked being challenged or stretched by a topic. For example, 28% of JUMP pupils interviewed in May mentioned being able to complete difficult tasks as a positive aspect of learning mathematics, but only 5% of IMPACT pupils did so.

Pupils in both groups also mentioned that they liked that mathematics was important for everyday life. This was particularly true of the second set of interviews, where the real-life relevance of mathematics was mentioned by 18% of JUMP and 13% of IMPACT pupils. Although not cited by any pupil in the first interviews, 10% of IMPACT pupils interviewed in May said that there were *no* best things about mathematics lessons.

The most commonly cited "worst thing" was that mathematics could be too hard or that a pupil was not able to answer the questions in class. Thirty-one percent of IMPACT pupils listed this as their worst thing in December, while 21% of JUMP pupils cited it in May. Mathematics being boring or doing too much repetitive work in class was another common negative aspect of mathematics, cited by 18% of JUMP pupils during the December interviews and 15% in May. While 15% of IMPACT pupils also mentioned the repetitive or boring aspect of mathematics during the December interview, only 3% did so at the end of the school year. Homework was not mentioned by any pupils in December, but was mentioned by 29% of

IMPACT pupils in May. Interestingly, one pupil – in a JUMP class – mentioned having to revise for "the Drumcondras" (i.e., the achievement tests used as part of the evaluation) as the worst thing about mathematics. Finally, at least 10% in each set of interviews indicated that there were no "worst things" about mathematics lessons.

Table 5.6: Percentages of pupils who identified various factors as the best/worst thing about learning mathematics

		December 2013		May 2014	
		JUMP (n = 39)	IMPACT (n = 39)	JUMP (n = 39)	IMPACT (n = 38)
Best	Specific topic	48.7	51.3	23.1	50.0
	Fun	17.9	5.1	7.7	2.6
	Games	2.6	20.5	2.6	2.6
	Important for life	5.1	5.1	17.9	13.2
	Being 'stretched'	12.8	7.7	28.2	5.3
	Are no best things	0.0	0.0	0.0	10.5
Worst	Specific topic	51.3	25.6	28.9	39.5
	Boring/repetitive	17.9	15.4	13.2	2.6
	Too hard/not being able to answer	15.4	30.8	21.1	18.4
	Homework	0.0	0.0	7.7	28.9
	Are no worst things	12.8	20.5	15.4	10.5

Being asked questions in class

Interviewers asked pupils if their teacher asked questions in class (all did, as expected) and followed up by asking how they felt when they were asked a question. In December 2013, a large majority of JUMP pupils (79%) indicated it was a wholly positive experience for them while the remaining 21% expressed mixed feelings. None felt it was a wholly negative experience (Table 5.7). By May 2014, only 45% indicated wholly positive feelings, 42% expressed mixed feelings, and 13% felt it was a wholly negative experience. In contrast, on both occasions slightly more than half of IMPACT pupils interviewed felt positive about being asked questions.

Table 5.7: Percentage of pupils indicating how they felt about being asked questions in their mathematics class

	December 2013		May 2014	
	JUMP (n = 38)	IMPACT (n = 39)	JUMP (n = 39)	IMPACT (n = 39)
Positive	78.9	56.4	44.7	53.8
Mixed	21.1	38.5	42.1	33.3
Negative	0.0	5.1	13.2	12.8

Examples of positive responses include "*Yes, cos I like telling the others how I did it*" and "*Yes, I really want to say it out and when I know it I feel better*". Ambivalent responses tended to focus on the distinction between the good feeling of knowing the correct answer and nervousness that it might be wrong: "*It's good if you know it. Otherwise, oh dear. But you have to try*" and "*I like questions when I know the answer but it's hard with everyone listening if I don't know it*". Negative responses focused largely on nervousness (sometimes mentioning additional pressure from the observer's presence): "*I get worried about getting things wrong,* *specially today* [being recorded]" or "*Not really, I'd be afraid I'd be wrong*". A number of pupils who expressed discomfort about being asked questions in front of the class mentioned that they preferred writing answers in their copybook or workbook.

Experience of mathematics instruction

To gauge pupils' perceptions of what activities typically occurred in mathematics lessons, the Pupil Questionnaire listed six teacher behaviours and asked pupils if these happened in their mathematics lessons. Table 5.8 summarises responses, showing the percentages of pupils who *agreed a lot* or *agreed a little* with the statements presented. Almost all pupils agreed that their teacher always explained what to do and asked if pupils understood the lesson. Between September and May, the percentage of pupils in JUMP classes agreeing with both statements increased by just over 5%, slightly higher than the 1% to 3% increase amongst IMPACT pupils. Thus, by May, 97% of JUMP pupils felt that their teacher always explained what they are expected to do and 95% agreed they always asked if pupils understood.

Roughly four in five pupils agreed that their teacher gave them fun things to do in maths lesson, with little difference by programme or by time of questionnaire completion. At the time of the first questionnaire, 80% of pupils in JUMP classes but only 68% of those in IMPACT classes agreed that their teacher let them play games in maths lessons. However, by the second questionnaire, just over three-quarters of pupils in each group agreed that they were let play games.

Table 5.8: Percentages of pupils who *agreed a lot/little* that various activities occurred during mathematics lessons, by time and programme

	September 2013		May 2014	
	% JUMP (N=271)	% IMPACT (N=238)	% JUMP (N=271)	% IMPACT (N=238)
My teacher always explains what we are expected to do	91.4	89.2	96.7	90.3
My teacher always asks do we understand stuff	89.8	87.2	95.2	90.3
My teacher often praises me	80.1	73.4	75.2	73.7
My teacher gets me to practice lots of examples	68.4	69.1	72.7	68.1
My teacher gives us fun things to do	78.9	78.4	80.4	78.4
My teacher lets us play games	79.8	67.5	77.2	76.4

Pupils were also asked about repeated practising of examples. In September 2013, similar percentages in each group agreed that this happened in maths lessons. By May 2014, an additional 4% of JUMP pupils and 1% fewer IMPACT pupils agreed that their teacher got them to practice lots of examples. Pupils were also asked if their teacher often praised them. IMPACT data remained quite static, with 73-74% agreeing this was true. However, there was a 5% drop in the number of JUMP pupils saying that their teacher often praised them.

Instruction during the observed lessons

During the course of pupil interviews, pupils were asked if the mathematics lesson just completed was similar to their normal lessons. By far the most common response (about two-thirds of pupils) was that the lesson was similar to their usual mathematics lesson, or similar apart from the presence of a camera. The next most common response was that the observed lesson involved more games or activities than a typical lesson: "*Usually there're games just once in a while, not all the time*". During the first set of observations, 16% of pupils mentioned that the lesson had more games than normal, compared to 9% after the second observations.

About 6% of pupils indicated that different materials might be used in other lessons or that the lesson varied by the day of the week: "S*ometimes we use workbooks and sheets and base 10 blocks and some materials*" / "*On Fridays we do stations*". A wide variety of other comments were made by pupils, most of which were very specific to the lesson topic. During the first set of interviews, four pupils also commented that the lesson was easy because they had been doing it for a while – although it is unclear if they meant the general topic or if the lesson itself had been practiced.

Use of learning strategies

Tables 5.9 and 5.10 show data from the Pupil Questionnaire, outlining pupil responses to how often they used each of a variety of learning strategies. Since the September 2013 baseline data from pupils were gathered shortly after the initial professional development on respective programmes, some early programme effects may be apparent in the first table.

At the start of the year, the strategy most commonly used by pupils in both groups was to do a sum in their head (83% of JUMP and 72% of IMPACT pupils indicated that they did so in every class or most classes, with only 7% [JUMP] to 10% [IMPACT] reporting that they hardly ever did so). By May 2014, the pupils in each programme who did a sum in their head in every lesson dropped to just over 65% (JUMP) and 59% (IMPACT), with 13-15% hardly ever doing so.

The strategy of trying to understand new material by drawing on pre-existing knowledge was popular in both programmes and at both time points. In the September and May responses, roughly three-quarters of pupils indicated that they used the strategy in every class or most classes. To gauge use of repeat procedures and memorisation (key to JUMP), pupils were asked about learning by heart. Regular use of memorisation increased only marginally amongst JUMP pupils (from 74% to 77% doing so in most or all classes), with a larger increase evident amongst IMPACT pupils (from 70% to 79%). In a related vein, JUMP pupils might have been expected to show increased use of repetition of examples as a strategy. However, the percentages of JUMP pupils who went through repeated examples in most or all lessons were largely unchanged from September (57%) to May (58%). Overall, IMPACT pupils used repetition of examples at much the same frequency as JUMP pupils.

Solving problems with classmates and considering multiple solutions might be considered characteristics of adherence to IMPACT. At the outset, problem-solving with classmates was reasonably common in both groups – 47% of JUMP and 53% of IMPACT pupils reported doing so in most or all classes. However, by May, the percentages of pupils regularly solving problems with their classmates had dropped to 33% of JUMP and 45% of IMPACT pupils. For thinking of more than one way to solve a problem, the percentage of IMPACT pupils who did so regularly rose from 63% in September to 69% in May. In contrast, the percentage of JUMP pupils doing so dropped from almost 70% to 65%.

There were few notable changes in the extent to which pupils thought about using maths in everyday life, apart from an increase (from 18% to 25%) in the number of IMPACT pupils who hardly ever did so. The second administration of the Pupil Questionnaire in May 2014 included one additional question on learning strategies. 73% of JUMP pupils and 64% of IMPACT pupils reported working on their own on a problem in most or all classes, slightly lower than might be expected given JUMP's emphasis on pupils working alone on examples in their workbooks.

Table 5.9: Percentages of pupils indicating various frequencies with which they engaged in learning strategies in mathematics lessons, **September 2013** responses

		Every class	Most classes	Some classes	Hardly ever
I work out a sum in my head	JUMP (N=263)	59.3	23.6	9.9	7.2
	IMPACT (N=229)	50.7	21.4	17.5	10.5
I try to understand new stuff by thinking about what I already know	JUMP (N=263)	54.4	20.9	12.5	12.2
	IMPACT (N=223)	40.4	33.2	17.0	9.4
When we do new things, I learn as much as I can by heart	JUMP (N=258)	48.8	25.6	14.3	11.2
	IMPACT (N=225)	40.9	29.3	21.8	8.0
I think of more than one way to get the answer to a problem	JUMP (N=262)	36.3	33.2	17.9	12.6
	IMPACT (N=226)	32.7	30.1	24.3	12.8
I go through examples again and again to help me remember them	JUMP (N=257)	34.6	22.6	21.4	21.4
	IMPACT (N=221)	36.7	21.7	25.3	16.3
I think about how I can use maths in everyday life	JUMP (N=268)	28.7	28.4	26.1	16.8
	IMPACT (N=228)	25.9	25.9	30.3	18.0
I work with my classmates to solve a problem	JUMP (N=258)	31.0	16.3	25.2	27.5
	IMPACT (N=225)	25.3	28.0	28.0	18.7

Table 5.10: Percentages of pupils indicating various frequencies with which they engaged in learning strategies in mathematics lessons, **May 2014** responses

		Every class	Most classes	Some classes	Hardly ever
I work out a sum in my head	JUMP (N=267)	33.7	31.8	21.0	13.5
	IMPACT (N=233)	36.5	22.7	25.8	15.0
I try to understand new stuff by thinking about what I already know	JUMP (N=267)	41.2	36.0	16.9	6.0
	IMPACT (N=234)	39.7	36.3	19.2	4.7
When we do new things, I learn as much as I can by heart	JUMP (N=262)	53.8	23.3	16.0	6.9
	IMPACT (N=234)	44.0	35.0	15.4	5.6
I think of more than one way to get the answer to a problem	JUMP (N=268)	28.7	36.2	28.4	6.7
	IMPACT (N=234)	32.9	35.9	21.4	9.8
I go through examples again and again to help me remember them	JUMP (N=265)	29.4	28.7	25.3	16.6
	IMPACT (N=234)	25.2	28.2	27.8	18.8
I think about how I can use maths in everyday life	JUMP (N=268)	28.7	25.4	30.2	15.7
	IMPACT (N=236)	18.2	29.2	27.5	25.0
I work with my classmates to solve a problem	JUMP (N=264)	13.6	19.3	43.6	23.5
	IMPACT (N=229)	17.0	27.9	39.3	15.7
I work on a problem on my own	JUMP (N=269)	37.2	35.7	22.3	4.8
	IMPACT (N=237)	26.6	35.9	25.7	11.8

Summary

Many of the attitudes expressed by pupils varied little by programme or by time. For example, a large minority of pupils in both programmes indicated that they did not like school, while almost half said they wished they did not have to study mathematics. Boys were far more likely than were girls to dislike school (and slightly more likely to dislike mathematics). There were slight increases in the percentages of pupils in both programmes agreeing that they learned interesting things in mathematics classes and that everyone could be good at mathematics, but a marked drop in the percentage of pupils interviewed who said that mathematics was their favourite subject. Questionnaire responses suggested that, in JUMP, boys' attitudes to mathematics improved more than girls' attitudes during the evaluation, while the reverse was true for IMPACT. However, these changes were small.

Pupils were asked about the activities that they or their teacher performed in a typical mathematics class. Again, there were very few differences or changes apparent. Irrespective of time of year or programme, most pupils said that their teacher always explained what to do and checked their understanding. Repeated practising of examples was equally common in each programme in September, but slightly more common in JUMP classes in May. JUMP pupils were also more likely to report regularly working on a problem on their own in class.

In contrast to questionnaire data, the interviews did reveal some differences by programme. Pupils in JUMP classes were far more likely to mention (unprompted) that they enjoyed being challenged or stretched in their mathematics lessons. Also, JUMP pupils were overwhelmingly positive about enjoying mathematics and preferring their current mathematics lessons to those in Second class, whereas the IMPACT pupils interviewed in May were far less positive than those interviewed earlier in the year. However, it should be borne in mind that IMPACT pupils' views probably relate to a mixture of IMPACT and non-IMPACT maths lessons, since the manuals covered material from just two strands and many teachers did not treat the programme principles as transferable across strands.

Chapter 6
Teachers

As with pupils, teachers' attitudes were assessed using start- and end-of-year questionnaires and two midpoint interviews. Teacher interviews were conducted immediately after classroom observations. This chapter presents the resultant data under three main headings, the first of which provides some information about participating teachers (demographic information, professional development, and confidence in relation to teaching mathematics). The second section describes teachers' accounts of typical mathematics lessons in each programme, including lesson length, materials and strategies used, and grouping practices. The final section describes teacher views on their assigned programme and the evaluation in general.

Given the timing of the questionnaires and interviews, it would be expected that the initial questionnaires would show few differences between teachers in each programme (as they were at that stage aware of their assigned programme, but almost all completed the questionnaire before attending any related training). However, as the second questionnaire was delivered towards the end of the school year, it would be expected to show marked differences between the teaching practices of teachers in each programme.

In five classes, teacher questionnaire and interview information was available for *two* teachers (in situations where a resource teacher spent a significant amount of time jointly teaching the class with the regular class teacher). In such cases, some information was aggregated to the class level. For example, if reported teaching experience was six and 10 years, then the class was assigned a notional aggregate of eight years. Other information (e.g., gender) did not lend itself to aggregation and is reported for all responding teachers.

Teacher characteristics

The 27 classes were taught by a total of 32 teachers, of whom 28 were female and four male. All classes taught by two teachers had two female teachers. The mean teaching experience for those teaching JUMP classes was 19.4 years, considerably higher than the mean of 11.5 years for IMPACT classes. The difference was largely attributable to three JUMP classes which were taught by teachers with more than 30 years of experience, whereas the most experienced IMPACT teacher had 21 years' experience. Generally, classes in both groups were taught by experienced teachers, and only one (IMPACT) class was taught by a teacher with less than three years of teaching experience.

Professional development

As part of the Teacher Questionnaire administered in September 2013, teachers were asked how many days of mathematics-related Continuing Professional Development (CPD) they had attended over the previous three years, exclusive of CPD related to the present evaluation. Responses ranged from none to 10 days, with JUMP teachers attending an average of 3.4 days, compared to 2.2 for IMPACT teachers (Table 6.1). The second Teacher Questionnaire, administered in May 2014, asked about experience of CPD (again, other than that related to the evaluation) during the year. IMPACT teachers attended 1.2 such days during the year, compared to 0.6 of a day attended by JUMP teachers. In total, JUMP teachers attended 4 days of CPD up to the end of May 2014, compared to 3.4 days for IMPACT teachers.

Table 6.1: Mean number of days spent on mathematics-related CPD (external to the evaluation), prior to and during the evaluation

	JUMP (N=13)	IMPACT (N=14)
Maths-related CPD, Sept 2010 – 2013	3.4 days	2.2 days
Maths-related CPD, 2013/14 academic year	0.6 days	1.2 days
Total, Sept 2010 – May 2014	4.0 days	3.4 days

Programme-specific professional development

As already noted in Chapter 1, some programme-specific training was made available to participating teachers. An initial introduction to each programme was conducted in early September 2013, lasting for most of the day. Each session was recorded and made available online to participating teachers. Two subsequent CPD events were organised for each programme: a webinar in November 2013 and a webinar in February 2014. In addition, in early 2014 teachers were given contact details for other teachers in their programme, to facilitate discussion and sharing of practice.

Chapter 1 provides detail on the numbers who attended each CPD session. It is notable that six of 17 JUMP and six of 15 IMPACT teachers (counting both class and resource teachers) did not attend the initial CPD day. Such relatively low attendance levels are partly attributable to very short notice given, in turn partly attributable an unexpected delay in one element of project funding. However, of the non-attenders, five did not watch the online recording of the CPD day, nor did they participate in the November webinar, while one did not participate in the February webinar either. Thus, three IMPACT and two JUMP teachers spent half of the academic year without any training for their assigned programme.

Teacher confidence and preparation

Teachers were asked about the degree to which they felt confident engaging in certain activities related to teaching mathematics, and how well prepared they felt to teach each of the curriculum strands. Table 6.2 shows the percentages of [all responding] teachers who felt they were *very well prepared* to teach strands,[1] with ratings more positive in the May than the September administration of the Teacher Questionnaire. However, this may not entirely reflect programme effects. Some of those who attended the initial CPD in September 2013 completed the Teacher Questionnaire shortly after they completed the MKTQ-S. As this is a reasonably difficult assessment of teacher knowledge for teaching mathematics, it may have had the unintended consequence of depressing some self-ratings. Also, it is worth repeating that IMPACT manuals did not cover Measures, Data, or Algebra, although the programme's principles are transferable across strands. It is thus less likely that changes in the confidence of IMPACT teachers on Measures, Data, and Algebra were programme-related.

Data was the strand with the lowest initial ratings. In September, only nine of 16 JUMP teachers (56%) and eight of 14 IMPACT teachers (57%) felt *well prepared* to teach Data, although by May 2014 almost all felt well prepared. Measures also had relatively low percentages of teachers who felt *well prepared* to teach it in September (56% of JUMP and 64% of IMPACT teachers). However, while most JUMP teachers felt well prepared for Measures in May 2014, almost one quarter of IMPACT teachers remained less than fully confident. In September, approximately four out of five teachers felt *well prepared* to teach Number, rising to

[1] With the exception of a single (IMPACT) teacher who reported feeling *not well prepared* to teach Algebra, all responses not shown in Table 6.2 are *somewhat prepared.*

87% of JUMP and 92% of IMPACT teachers in May. Separate ratings for Algebra were not requested in September, as it was expected that most teachers would treat Algebra as part of Number. However, as described in Chapter 3, our subsequent analysis of JUMP materials suggested that Algebra was more heavily represented in JUMP than in either the PSMC or Irish textbooks. Therefore, the May questionnaire was modified to include a separate rating for this strand, but JUMP and IMPACT teachers reported fairly similar levels of confidence (81% and 77%, respectively, felt *well prepared* to teach Algebra). The percentage of IMPACT teachers *well prepared* to teach Shape and Space was largely unchanged (from 86% to 85%), but the percentage of JUMP teachers who felt *well prepared* to teach it increased from 75% to 87%.

Table 6.2: Percentages of teachers indicating they felt *very well prepared* to teach various PSMC strands

	September 2013		**May 2014**	
Number	JUMP (N=16)	81.3	JUMP (N=16)	87.5
	IMPACT (N=14)	78.6	IMPACT (N=13)	92.3
Shape & Space	JUMP (N=16)	75.0	JUMP (N=16)	87.5
	IMPACT (N=14)	85.7	IMPACT (N=13)	84.6
Measures	JUMP (N=16)	56.3	JUMP (N=16)	87.5
	IMPACT (N=14)	64.3	IMPACT (N=13)	76.9
Data	JUMP (N=16)	56.3	JUMP (N=16)	87.5
	IMPACT (N=14)	57.1	IMPACT (N=13)	92.3
Algebra			JUMP (N=16)	81.3
			IMPACT (N=13)	76.9

In relation to specific teaching skills, most teachers initially expressed only medium confidence levels. Based on responses supplied in September 2013, a minority felt *very confident* in most of the six instructional activities listed (Table 6.3), and almost none indicated that they were *not confident* on any aspect.[2] Less than half of teachers in each programme initially felt *very confident* of their ability to provide challenging tasks for capable pupils, to adapt their teaching to engage pupils' interest, or to work with lower-achieving pupils. Further, less than half of JUMP teachers (and slightly more than half of IMPACT teachers) felt *very confident* connecting different mathematics topics, showing a variety of methods for doing calculations, and teaching real-life problem-solving.

By the second administration of the questionnaire in May 2014, self-ratings had increased for five of the six skills. There were increases in the number of JUMP teachers who were *very confident* in connecting mathematics topics, and in JUMP and IMPACT teachers who were *very confident* showing a variety of methods for doing calculations. For JUMP, the number of teachers who felt *very confident* working with lower-achieving pupils rose from five to nine. The exception was on ability to adapt teaching to engage pupils' interest. By May, only two JUMP teachers (down from six initially) reported being *very confident.*

[2] As with Table 6.2, only the most positive response option (*very confident*) is shown. The vast majority of responses not shown are in the middle (*somewhat)* response category.

Table 6.3: Percentages of teachers that felt *very confident* on various aspects of mathematics instruction

	September 2013		May 2014	
Connecting one mathematics topic to another	JUMP (N=16)	37.5	JUMP (N=16)	62.5
	IMPACT (N=15)	66.7	IMPACT (N=14)	78.6
Showing pupils a variety of methods for doing calculations	JUMP (N=16)	37.5	JUMP (N=16)	68.8
	IMPACT (N=15)	53.3	IMPACT (N=14)	78.6
Providing challenging tasks for capable pupils	JUMP (N=16)	37.5	JUMP (N=16)	43.8
	IMPACT (N=15)	33.3	IMPACT (N=14)	42.9
Adapting my teaching to engage pupils' interest	JUMP (N=16)	37.5	JUMP (N=16)	13.3
	IMPACT (N=15)	40.0	IMPACT (N=14)	57.1
Working with lower-achieving pupils	JUMP (N=16)	43.8	JUMP (N=16)	50.0
	IMPACT (N=15)	33.3	IMPACT (N=14)	64.3
Teaching real-life problem-solving	JUMP (N=16)	43.8	JUMP (N=16)	56.3
	IMPACT (N=15)	60.0	IMPACT (N=14)	71.4

Collaboration with other teachers

Teachers were asked how often they discussed how to teach a particular topic with other teachers, and how often they worked with other teachers to try out new ideas. At the start of the year, the most common response from teachers in both groups (50% in JUMP and 60% in IMPACT) was that they discussed how to teach a particular topic with other teachers about two to three times a month (Table 6.4). However, 13% of JUMP and 27% of IMPACT teachers reported never or almost never discussing how to teach a topic.

There was a slight decrease in discussions with colleagues over the duration of the evaluation. May responses show that one quarter of teachers in the JUMP programme and over one third of those in IMPACT reported never or almost never discussing the teaching of particular topics with other teachers. Further, none of the IMPACT teachers had daily or almost daily discussions with colleagues, although the number of JUMP teachers engaging in daily discussions with colleagues increased by 6% (i.e., one teacher).

Table 6.4: Percentages of teachers indicating how often they discussed the teaching of particular topics with other teachers

	September 2013		May 2014	
	JUMP (N=16)	IMPACT (N=15)	JUMP (N=16)	IMPACT (N=14)
Daily or almost daily	18.8	6.7	25.0	0.0
1-3 times a week	18.8	6.7	12.5	14.3
2-3 times a month	50.0	60.0	37.5	50.0
Never or almost never	12.5	26.7	25.0	35.7

Table 6.5 shows the frequency with which teachers worked with other teachers to try out new ideas. As can be seen, responses from teachers in the JUMP programme varied little across the two time periods, while there was a very slight increase in frequency of working together amongst teachers in IMPACT. Generally, though, less than half of teachers in either programme or at either time point worked together with other teachers to try out new idea on a very regular basis.

Table 6.5: Percentages of teachers indicating how often they worked together with other teachers to try out new ideas

	September 2013		May 2014	
	JUMP (N=16)	IMPACT (N=15)	JUMP (N=16)	IMPACT (N=14)
Daily or almost daily	12.5	6.7	12.5	0.0
1-3 times a week	25.0	6.7	31.3	21.4
2-3 times a month	25.0	46.7	18.8	42.9
Never or almost never	37.5	40.0	37.5	35.7

Typical mathematics lessons

Teachers were asked a number of questions to establish what happened during mathematics lessons, including lesson duration, materials used, grouping practices, and general class activities.

Lesson length

In September, the mean time spent on mathematics by JUMP classes was 286 minutes per week, slightly higher than the mean of 277 in IMPACT classes (Table 6.6). By May 2014, the difference was reversed and pupils in IMPACT classes were spending slightly more time in mathematics lessons than were JUMP pupils (291 versus 282 minutes per week). Thus, in May the average IMPACT lesson lasted almost an hour (58 minutes) while the average JUMP lesson was 56 minutes.

Table 6.6: Mean number of minutes spent teaching mathematics in participating classes, per week

	JUMP (N=13)	IMPACT (N=14)
September 2013	286.1	276.9
May 2014	282.0	291.5

Materials Used

Teachers were asked about the frequency (*most or all lessons*, *once or twice a week*, *once or twice a month*, and *rarely or never)* with which they used various materials in their mathematics lessons. Tables 6.7 summarises some responses, showing the percentages of teachers indicating that they used materials at least once or twice a week (i.e., combining the two most frequent options). As might be expected, real-life materials were a common feature of IMPACT lessons, featuring at least weekly in lessons at the start and towards the end of the school year. However, all but two of the 16 JUMP teachers also reported at least weekly use of real-life materials.

Weekly use of manipulatives was slightly more common in IMPACT classes, especially in May, when 75% of JUMP and 100% of IMPACT teachers reported their use. Regular use of games was common across both programmes, while tablebooks were a more frequent feature of IMPACT than JUMP lessons. In September, slightly less than half of JUMP teachers used tablebooks (compared to almost all IMPACT teachers), rising to just over half (56%) by May.

In sum, with the exception of two additional teachers regularly using tablebooks, there was little change in the use of these materials by teachers in the JUMP programme over the course of the year.

Table 6.7: Percentages of teachers reporting that they used various materials at least once or twice a week in mathematics lessons

	September 2013		May 2014	
	JUMP (N=16)	IMPACT (N=15)	JUMP (N=16)	IMPACT (N=15)
Real-life materials	87.5	100.0	87.5	100.0
Manipulatives	81.3	92.9	75.0	100.0
Games	87.5	80.0	87.5	92.3
Tablebooks	43.8	93.3	56.3	66.7

As regular use of the pupil workbook, along with supplementary pupil materials if appropriate, might be considered an important aspect of adherence to JUMP, Table 6.8 shows the detail of teacher responses about using such materials. At the start of the year, almost every teacher indicated that they would normally use textbooks in most or all lessons with Third class pupils. However, by May, almost one-third of JUMP teachers (and almost one half of IMPACT teachers) indicated that they were using textbooks[3] only once or twice a week.

There was also a noticeable change in the frequency of use of workbooks/worksheets. In September, 25% of JUMP teachers indicated that they generally used worksheets in most or all lessons (at that stage the JUMP pupil materials had not been delivered to schools). By May, 69% reported using workbooks/worksheets in most or all lessons, yet a sizeable minority of teachers (31%) still did not use them in most lessons. In contrast, the percentage of IMPACT teachers who regularly used workbooks/worksheets dropped slightly.

Table 6.8: Frequency with which teachers reported using any type of textbooks and workbooks/worksheets

		September 2013		May 2014	
		JUMP (N=16)	IMPACT (N=15)	JUMP (N=16)	IMPACT (N=15)
Textbooks	Most or all lessons	93.8	93.3	68.8	53.8
	Once or twice a week	6.3	6.7	31.3	46.2
	Once or twice a month	0.0	0.0	0.0	0.0
	Rarely or never	0.0	0.0	0.0	0.0
Workbooks/ worksheets	Most or all lessons	25.0	46.7	68.8	33.3
	Once or twice a week	62.5	46.7	25.0	41.7
	Once or twice a month	6.3	6.7	0.0	16.7
	Rarely or never	6.3	0.0	6.3	8.3

Teaching strategies

Teachers were asked how often they used a variety of teaching strategies during their mathematics lessons, including breaking ideas down into very simple steps, asking pupils what they learned, relating lessons to everyday life, bringing interesting materials to class, and teaching how to solve a problem by using multiple similar problems.

In September 2013, 94% of JUMP and 67% of IMPACT teachers said that they broke ideas down into very small steps in every or almost every lesson (Table 6.9). This approach is a key component of JUMP, but the difference is not attributable to JUMP itself, as the questionnaires were largely completed before teachers attended initial training. When Teacher

[3] Generally, textbooks were understood to mean Irish textbooks, not the JUMP workbooks. However, some teachers may have counted the JUMP workbooks in the "textbook" category at the end of the year.

Questionnaires were next administered in May, breaking down ideas into smaller steps was a more common feature of IMPACT than of JUMP lessons, despite the importance of sequential, scaffolded steps featuring very prominently in JUMP materials.

Table 6.9: Percentages of teachers reporting the frequency with which they break ideas down into very small steps during mathematics lessons

	September 2013		May 2014	
	JUMP (N=16)	IMPACT (N=15)	JUMP (N=16)	IMPACT (N=13)
Every/almost every lesson	93.8	66.7	81.3	92.3
About half the lessons	0.0	33.3	18.8	7.7
Some lessons	6.3	0.0	0.0	0.0
Never	0.0	0.0	0.0	0.0

At the start of the year, only two teachers in each programme (12% and14%) asked pupils what they had learned in most or all classes (Table 6.10). However, May responses indicate that teachers in both programmes, but particularly in IMPACT, increased the frequency with which they checked pupil understanding of lessons. As a corollary, the number of JUMP teachers who only checked intermittently, if at all, fell from 37% to 19%, while the number of such IMPACT teachers fell from 57% to 14% (i.e., from eight to only two teachers).

Table 6.10: Percentages of teachers reporting the frequency with which they asked pupils what they had learned after mathematics lessons

	September 2013		May 2014	
	JUMP (N=16)	IMPACT (N=15)	JUMP (N=16)	IMPACT (N=14)
Every/almost every lesson	12.5	14.3	31.3	50.0
About half the lessons	50.0	28.6	50.0	35.7
Some lessons	37.5	35.7	18.8	14.3
Never	0.0	21.4	0.0	0.0

There was also an increase over the school year in the frequency with which mathematics lessons were related to everyday life. In September, a little less than half of teachers in each programme related mathematics lessons to daily life on a daily or almost daily basis (Table 6.11). By May, the percentages doing so had risen to 63% of JUMP and 71% of IMPACT teachers. Nonetheless, two JUMP teachers (13%) linked real life to less than half of their lessons.

Table 6.11: Percentages of teachers reporting the frequency with which they related mathematics lessons to daily life

	September 2013		May 2014	
	JUMP (N=16)	IMPACT (N=15)	JUMP (N=16)	IMPACT (N=14)
Every/almost every lesson	43.8	40.0	62.5	71.4
About half the lessons	37.5	26.7	25.0	28.6
Some lessons	18.8	33.3	12.5	0.0
Never	0.0	0.0	0.0	0.0

In September, a large minority of teachers in each programme (31% in JUMP and 43% in IMPACT) indicated that teaching problem-solving by using multiple similar problems was a feature of all or almost all their lessons (Table 6.12). However, a large minority also indicated that they only did so in some lessons. Amongst JUMP teachers, the numbers using multiple similar problems in most/all lessons doubled between September and May, while the number who did so in only some lessons almost halved. IMPACT teachers also showed an increase in use of teaching problem-solving by using multiple similar problems, but not to the same extent as JUMP.

Table 6.12: Percentages of teachers reporting the frequency with which they taught problem-solving using multiple similar problems

	September 2013		May 2014	
	JUMP (N=16)	IMPACT (N=15)	JUMP (N=16)	IMPACT (N=14)
Every/almost every lesson	31.3	42.9	62.5	42.9
About half the lessons	25.0	21.4	12.5	42.9
Some lessons	43.8	35.7	25.0	14.3
Never	0.0	0.0	0.0	0.0

In most cases, teachers reported an increase in the use of a particular strategy over the period of the evaluation. An exception was JUMP teachers bringing "interesting materials" to class. In September, almost one third of JUMP teachers brought interesting materials to class for most or all lessons, but this fell to 6% (one teacher) by May (Table 6.13). In contrast, IMPACT teachers increased the frequency with which they brought materials into lessons, although one teacher did indicate that interesting materials were *never* a feature of their mathematics classes.

Table 6.13: Percentages of teachers reporting the frequency with which they brought interesting materials to class

	September 2013		May 2014	
	JUMP (N=16)	IMPACT (N=15)	JUMP (N=16)	IMPACT (N=14)
Every/almost every lesson	31.3	6.7	6.3	14.3
About half the lessons	37.5	26.7	43.8	57.1
Some lessons	31.3	66.7	50.0	21.4
Never	0.0	0.0	0.0	7.1

Grouping practices

The most common grouping practice at the start of the year was whole class teaching (used in *most lessons* in 93% of IMPACT and 60% of JUMP classes) (Table 6.14). Among teachers in the JUMP programme, the extent to which whole class teaching was used changed little from September to May. Similarly, at both the start and end of the year, about one third of JUMP teachers and one fifth of IMPACT teachers reported that pair work was used in most lessons. In contrast, the extent of small group work changed during the year. In JUMP classes, small group work was slightly less frequently used at the end of the year (12% rarely or never used it), while the percentage of IMPACT teachers using small group work in most lessons increased from 7% to 43%.

In September, 53% of JUMP and 80% of IMPACT classes featured individual work in most lessons. However, by May, the extent to which individual work was used in lessons was almost identical across both programmes (69% of teachers in each case used individual work in most lessons). Thus, individual work was slightly less common than might be expected in JUMP classes and slightly more common in IMPACT classes.

Table 6.14: Percentages of teachers reporting the frequency with which they used various grouping practices

		September 2013		May 2014	
		JUMP (N=15)	IMPACT (N=15)	JUMP (N=16)	IMPACT (N=15)
Whole class	Most lessons	60.0	93.3	62.5	76.9
	Some lessons	40.0	6.7	31.5	23.1
	Rarely or never	0.0	0.0	6.3	0.0
Small groups	Most lessons	60.0	6.7	50.0	42.9
	Some lessons	40.0	93.3	37.5	57.1
	Rarely or never	0.0	0.0	12.5	0.0
Pairs	Most lessons	33.3	20.0	31.3	21.4
	Some lessons	66.7	73.3	62.5	78.6
	Rarely or never	0.0	6.7	0.0	0.0
Individual	Most lessons	53.3	80.0	68.8	69.2
	Some lessons	46.7	13.3	31.3	30.8
	Rarely or never	0.0	6.7	0.0	0.0

Pupil activities

Teachers were asked how often they asked pupils to do the following: listen to the teacher explain how to solve problems; memorise rules, procedures and facts; work problems individually or with peers, with the teacher's guidance; work problems together as a class, with the teacher's guidance; explain their answers; and self-assess their mathematical performance.

Response options were *every/almost every lesson, about half the lessons, some lessons* and *never*. For ease of presentation, Table 6.15 shows the percentages of teachers who got pupils to engage in an activity in most or at least half of their lessons. Differences by programme are not any more pronounced in the May responses than in the initial responses in September. For example, May responses show that 100% of teachers in each programme asked pupils to explain their answers in at least half of lessons, and over four in five got pupils to work problems in small groups.

Table 6.15: Percentages of teachers indicating that they asked pupils to engage in various activities in at least half of their mathematics lessons

	September 2013		**May 2014**	
Listen to the teacher explain how to solve problems	JUMP (N=16)	56.3	JUMP (N=16)	62.5
	IMPACT (N=15)	75.0	IMPACT (N=14)	57.1
Memorise rules, procedures and facts	JUMP (N=16)	62.5	JUMP (N=15)	60.0
	IMPACT (N=15)	40.0	IMPACT (N=14)	50.0
Work problems alone			JUMP (N=16)	87.5
			IMPACT (N=14)	78.6
Work problems in a small group			JUMP (N=16)	87.5
			IMPACT (N=14)	85.7
Work problems together as a whole class	JUMP (N=16)	75.0	JUMP (N=16)	87.5
	IMPACT (N=15)	73.3	IMPACT (N=14)	78.6
Explain their answers	JUMP (N=16)	81.3	JUMP (N=16)	100.0
	IMPACT (N=15)	66.7	IMPACT (N=14)	100.0
Self-assess their performance	JUMP (N=16)	56.3	JUMP (N=16)	75.0
	IMPACT (N=15)	26.7	IMPACT (N=14)	42.9

Asking pupils to work alone on problems in at least half of lessons was very common in both programmes (79% in IMPACT and 87% in JUMP) but universal in neither. Regular memorisation of rules, procedures and facts was a feature of only 60% of JUMP classes in May, largely unchanged since September, and not markedly different from IMPACT (50% of classes). Asking pupils to self-assess their performance was one of the few areas in which there were differences by programme. Three-quarters of JUMP teachers did so in at least half of classes, compared to 43% of IMPACT teachers. However, the September responses show a pre-existing difference (56% in JUMP versus 27% in IMPACT), suggesting it may be unrelated to programme effects.

Teacher views on programme/evaluation

All teachers were interviewed at two points during the year, coinciding with the first and second set of classroom observations. Interviews took place immediately after the observations and covered topics such as level of comfort in using the assigned programme in classes, views on the efficacy of the programme, and views on the evaluation itself. This section summarises teacher views on the CPD provided to participants, their degree of comfort in using the respective programmes, and their views on the general efficacy of the programme and the implementation of the evaluation.

Much of the preceding chapters compares JUMP and IMPACT, to establish initial similarities, and identify any eventual differences. The following sections have a much stronger focus on JUMP itself, as the focus is on aspects of the assigned programme that teachers perceived as helpful or problematic, and on the extent to which they felt that programme practices and methodology were a feature of their classrooms.

Satisfaction with and uptake of CPD

As noted in Chapter 1, teachers received very little notice for the initial training day for their respective programmes (partly due to a delay in release of some project funding, which in turn contributed to difficulty in scheduling Dr Mighton's availability). This led to poor attendance levels. Of those that did attend, satisfaction was higher among IMPACT than JUMP teachers. Of the nine class teachers who attended the initial JUMP CPD, only three expressed wholly positive views on the quality of the training, four expressed mixed views, and two were generally negative. For IMPACT, seven of the eight class teachers who attended were wholly positive and the remaining teacher expressed ambivalent views. IMPACT was commended for the practical nature of the session. The main criticisms of the JUMP session were that the training was not practical and that there was too much material.

> "*I thought it was very good. I liked the comparison of Canadian test results, the methods, the visual approach.*"

> *"It was not what I thought. It wasn't training – it was research oriented."*

> *"It was very academic – not very practical. Like, by the end of training, we still hadn't looked at the programme books.*"

> *"It was good – teachers have discretion about how to use it. Lot of material though.*"

Attendance was also relatively poor for the November 2013 webinars. Six of the eight JUMP attendees were positive and two were negative in their comments. Three IMPACT attendees expressed positive comments, three expressed negative comments and two offered no opinion. A few complaints related to the format and organisation of the training (face-to-face, and more notice preferred), while positives included being able to raise questions, and talking to other teachers.

Attendance at the February webinar was slightly improved due to a longer notice period. While none of the attendees at either session expressed wholly negative views on the webinars, many were unhappy with the format and with some technical problems that arose during the sessions. Seven JUMP and five IMPACT attendees specifically criticised the webinar format. For example:

"*John Mighton's enthusiasm was infectious but the technology was not satisfactory.*"

"*We signed up but there was some sort of system overload and we were asked to sign out.*"[4]

As a number of teachers had referred to communicating with other teachers about the programmes as a major advantage of CPD, in early 2014 a teacher email network was offered to those interested, as was access to a discussion website for teachers. Five JUMP and two IMPACT teachers availed of the opportunity to contact other teachers directly while six JUMP and two IMPACT teachers had looked at the websites by May 2014.

Generally, the extent of engagement in CPD by teachers in each programme was lower than might be hoped for, and many teachers expressed dissatisfaction that short notice and poor communication had prevented them from attending.

"*We needed more notice for the* [initial] *training* [which she could not attend]. *It's very daunting when you are in the school and the books arrived without the manual.*"

"*You need more training and support. You can be isolated and get worried that you are not covering the course properly or won't get it finished*"

That aside, relatively few availed of chances to look at the programme websites, view the training videos, or contact other teachers.

Extent to which programme was used

To gauge how much the programme featured in day-to-day classes, teachers were asked to rate their own level of comfort with using their respective programme, to indicate if they used the programme in conjunction with any other programmes, and how much of each PSMC strand they covered using their assigned programme. Specific to JUMP, teachers were asked about their use of the JUMP Confidence Building Unit.

Confidence-building unit (CBU)

As noted in Chapter 3, the CBU is a feature of JUMP that is meant to be used for up to two weeks, at the start of the school year. It covers fractions and aims to promote confidence by helping pupils master *procedures* usually tackled by pupils at more advanced grade levels, but does not aim to teach the topic in conceptual depth. Of the seven JUMP classes that used the CBU, six used it at the start of the year and one used it around Christmas. Three teachers indicated that they did not use it at all while three gave it to colleagues for use with older classes.

The length of time for which the CBU was used varied considerably – ranging from one week to "*nearly until Christmas*". Most of the teachers interviewed seemed unsure for how long

[4] The number of teachers who attempted to login to the webinar exceeded the number who had formally signed up to take part, resulting in access difficulties for some.

they were supposed to use the CBU, and those who used it for a longer time period expressed dissatisfaction.[5]

"*We did it 'til Halloween. It was too long. I wouldn't make that mistake again and we're now playing catch-up.*"

"We didn't finish it as we had to start the first book. I thought the pacing and time guidelines were poor."

Views on the benefits of the CBU were mixed, with some (non-users) expressing the view that it was too difficult for the target grade. Three teachers thought it was worthwhile, and another noted that while the children enjoyed it, she was unsure of the benefits. One teacher (who used the CBU as instructed for 10 days at the start of the year) commented that "*It became quite complicated after a while. I don't understand the theory of having something so difficult as a confidence-building exercise.*"

Combining with other programmes

As part of the first interview, teachers were asked if their school was involved in any other mathematics projects, and if they used methods or materials from other sources. Just over half (54%) of JUMP teachers said that they tended to just use JUMP. The remainder (46%) were using JUMP in combination with another programme or resource – either the Maths Recovery programme or with their regular textbooks.

"I mix and match with Maths Recovery. I prefer the Maths Recovery training. It's three hours a week for five weeks."

Four (31%) IMPACT teachers said that they were only using IMPACT,[6] seven (54%) were using it in combination with another programme or resource, while two said that they were not using IMPACT at all.

Comfort level

Teachers were asked to assess their own skill or comfort level in using their assigned programme (choosing from *skilled*, *fairly comfortable*, *novice*, *non-user*, and *former user*). During the first set of interviews in December 2013/January 2014, all JUMP teachers described themselves as fairly comfortable using the programme (somewhat surprisingly, including one teacher interviewed in January 2014 who had not used the programme at all in 2013). Seven of the 13 IMPACT teachers also said they were fairly comfortable using IMPACT, three felt they were novice level users and one described herself as a skilled user. Two teachers indicated that they were not using IMPACT.

By the next set of interviews in May, eight JUMP teachers continued to describe themselves as fairly comfortable using JUMP, while four felt they were now skilled users. Four IMPACT teachers described themselves as skilled IMPACT users, seven as fairly comfortable users, while two continued to feel they were novice users. None of the teachers interviewed in May indicated that they were non-users.

[5] The teacher manual for the CBU *does* advise that teachers should spend no more than two weeks on the CBU, but the advice is somewhat buried in the text.

[6] As IMPACT manuals addressed only two PSMC strands, this presumably referred to using IMPACT *principles*, not materials, to cover Third class lesson content.

Strand coverage by programme

In the May interviews, teachers were asked how much of each of the strands they had covered during the year using their assigned programme. IMPACT responses were heavily skewed by the materials they had received, which focussed only on Shape and Space and on aspects of Number. Thus, a minority used IMPACT (principles) to teach Data, Algebra or Measures, while 92% used IMPACT to teach most or all of Shape and Space, with 54% using it to teach most or all of Number.

Table 6.16 shows the extent to which JUMP materials were used by teachers. For Number and Algebra, one teacher (different for each strand) did not use JUMP to cover the strand at all, while a sizeable minority (six teachers for Measures and five for Data) only used JUMP for some strand content. Algebra was the strand area most likely to be covered using JUMP, with 10 of the 13 JUMP teachers indicating that JUMP was used to cover most or all of the strand. This may be related to the finding reported in Chapter 3 that JUMP devotes slightly more attention to Algebra than does the PSMC. In sum, while used fairly widely, many teachers were selective in how and when they used JUMP, and many used it in a modified form.

> "*I'd usually read the JUMP unit then decide what to do. Measures didn't have enough on capacity and length, for example.*"
>
> "*I'd use it but I adapted many sections. Money for example.*"

Table 6.16: Number of JUMP teachers reporting how much of each strand was covered using JUMP materials or methods, May 2014

	None/almost none	Some of it	Most/all of it
Number	1	3	9
Shape and Space*	0	4	8
Measures	0	6	7
Data	0	5	8
Algebra	1	2	10

*One teacher (not included in this row) had not yet covered Shape and Space.

Programme efficacy

Views on programme efficacy were sought through questions about alignment with the PSMC, pupils' perceptions, how effective the programmes were for different types of pupils, and general thoughts on programme benefits and weaknesses.

Curriculum alignment

Teachers were twice asked to rate how well their assigned programme matched the PSMC on a scale of 1 ("not at all") to 6 ("perfect match"). Initial ratings for JUMP group ranged from 1 to 5, averaging 4.1 (equivalent to a reasonably good match to the curriculum) (Table 6.17). IMPACT ratings ranged from 2 to 6, averaging 4.0. This may be considered surprising for a programme designed explicitly to meet the objectives of the PSMC, but it is likely that some teachers lowered their ratings to account for the fact that IMPACT materials had not been developed for all strands. Five JUMP teachers criticised the "pitch" of JUMP materials – although some felt they were pitched too high and others too low. Another five critical comments related to the need to adapt some parts of the materials to the Irish context. The suitability of units about Canadian currency was cited as an issue.

In the second set of interviews, ratings for both programmes improved, with JUMP averaging 4.4 and IMPACT 4.9. Pitch of JUMP materials was raised as an issue by only one teacher, but differences in strands/topics covered were raised as an issue by five teachers. JUMP was criticised for over-emphasis on patterns, and insufficient emphasis on Number and problem-solving. Americanised spellings and terminology for money were also flagged as problems.

> *"Money, rivers, even the language is very Canadian."*

> *"Our system has more Number and more word problems. JUMP has an awful lot on patterns (it's too slow)."*

Table 6.17: Teacher ratings of the degree of match (scale 1 to 6) between programme and PSMC

	Dec/Jan		May	
	JUMP (N=13)	IMPACT (N=13)	JUMP (N=13)	IMPACT (N=13)
Mean	4.1	4.0	4.4	4.9
Range	1 – 5	2 – 6	3 – 5.5	4 – 6

Pupil engagement

Teachers were asked to rate their pupils' response to the assigned programme on a scale of 1 (extremely negative) to 6 (extremely positive). In December 2013/January 2014, teacher ratings were generally high across both groups, ranging from 3 – 6 for JUMP and from 4 – 6 for IMPACT. The mean rating in JUMP was 4.9, while in IMPACT it was 5.1. By May, IMPACT was rated slightly less positively (4.8) while JUMP retained an average rating of 4.9. Thus, teachers in both groups felt that their pupils were responding well to the programmes.

Particular aspects of JUMP to which teachers felt that pupils responded well were the hands-on approach (cited by three teachers in the second interviews) and the quality or quantity of JUMP materials (cited by five). While the lack of colour was raised as a criticism of JUMP workbooks, most teachers felt that the amount of content and the variety of activities for pupils were very positive aspects of the programme. In contrast, IMPACT was generally praised for allowing more talking about maths and less use of textbooks.

> *"They liked the hands-on bits, kept them going. It's [JUMP] a very busy programme – lots to be covered in each topic."*

> *"Without a doubt [pupils liked JUMP] – puzzles, Sudoku, Hanji very enjoyable."*

Target group

When first asked if they thought the programme was more beneficial for some pupils than others, seven JUMP teachers felt it was more effective for weaker pupils, five felt it worked for all ability levels and one was unsure of the target level. Of 11 teachers using IMPACT, six thought it worked best for weaker pupils and five viewed it as an all-ability programme.

> *"It's* [JUMP] *good for all. Brighter children can move quickly – extension exercises are good for these. The repetition is good for the weaker child. So it's well suited to all levels (and well written).*

By the second interviews, seven IMPACT teachers thought it worked well for all pupils, four thought it best for weaker and two for stronger pupils. Eight JUMP teachers thought JUMP was effective for all ability levels, three thought it best for weaker pupils, and one for stronger. In one JUMP class the two teachers involved disagreed – one believed JUMP worked best for weaker pupils while one believed it was most effective with stronger pupils.

"It challenged the 'bonus questions' group, but with weaker students there was no overall change" [teacher 1]. *"The breaking things down into small steps was very good for the weak student but it didn't challenge the stronger ones."* [teacher 2]

Generally, most teachers felt that all pupils gained from participation in their programmes.

"I thought it [JUMP] *was good for all. You can extend stronger pupils AND it definitely works for weaker pupils as it is so hands-on."*

Strengths and weaknesses

In December 2013/January 2014, 12 of 13 JUMP teachers agreed that there were some major benefits to their programme, with the remaining teacher undecided as to the merits of the programme at that stage. All 11 IMPACT teachers using the programme by January also identified programme benefits. In both groups, 10 teachers also identified major drawbacks to their assigned programme. By the May interviews, 12 JUMP teachers identified major benefits to using JUMP, while one felt it a somewhat useful programme, but not a major advance on normal classroom procedures.

Broadly, the number of aspects identified by JUMP teachers as major programme benefits decreased over the course of the two interviews while the number of IMPACT benefits identified increased. The benefits identified most often by JUMP teachers were that the programme helped build pupil confidence and positive attitudes to mathematics, and that the step-by-step approach was very good (seven teachers cited each in the first interviews, as did four in the second interviews).

"It breaks everything down, which is important. That builds up confidence in weaker pupils."

The quality and content of the JUMP pupil and/or teacher materials were praised by four teachers in the first interviews, rising to seven by May (when five IMPACT teachers also cited quality of materials as a major benefit).

"It [JUMP] *is well organised, planned, resourced. I like the way they introduce topics gradually – lots of examples, a confidence boost with simple examples. It covers a wide range of abilities, and has detailed teacher books."*

"The teacher book is great. Great to have an explanation and breakdown of how to teach a topic. That's not available with Irish textbooks. I like the focus on maths vocab too. "

Other positive aspects of JUMP identified by teachers included the repetition of language and procedures (five teachers in the first, and two in the second interviews) and the hands-on or practical nature of the programme (mentioned by two in the second interviews).

Almost all (all JUMP and 11 IMPACT teachers) also identified drawbacks to their assigned programme. For IMPACT, the main criticisms were related to pitch (with some feeling they had to revert to Second class material to get pupils up to speed) and specificity (the manual was perceived as a bit "waffly" and without enough practical lesson plans). For JUMP, the initial interviews revealed two major issues – the need for cultural adaptation and problems with the organisation of the project (each cited by eight teachers in the first interview and four in the second).

"*Canadian language and money."*

"You have to study the manual – some of the terminology is a bit strange. Things like multiples and digits."

"There were organisational mess-ups – the workbooks arrived without manuals. It made it daunting. I expected more hands-on resources as well as textbooks."

"The timeframe for covering the programme was unclear. We got the idea at the webinar, but not before then. The training needed more notice. It was very isolating – a lot of reading on your own that made it huge pressure at the start."

The nature of JUMP materials were criticised by six teachers in each set of interviews. Typical comments referred to the lack of colour and pictures in pupil materials or to the overall volume of teacher and pupil materials.

"The workbook is off-putting for children: lack of colour, lot of content on some pages. It suits some children and not others."

"I'd like to have had more resources, and for the manual to be better adapted to the curriculum."

"The manual is awkward (one massive folder) and wordy."

Some JUMP teachers criticised the amount of work expected from the teacher, but this was a more common complaint from IMPACT teachers (eight initially complained about the amount of extra work). Other complaints related to the relative coverage of strands and strand units. For example, one teacher complained that there was not enough emphasis on Measures in JUMP, one cited insufficient attention to the mechanics of subtraction and division, and another felt Number was not given in-depth treatment. In contrast, one teacher felt that too much time was spent covering money.

As a conclusion to the interview, 10 JUMP and eight IMPACT teachers offered general comments on their respective programmes. Four IMPACT teachers expressed an unqualified positive comment (e.g., "*It's wonderful for getting children to think and it has improved my own questioning*"), two wanted it to cover all strands, one wanted it more aligned to standardised tests, and one noted that it aligned closely with their (recent) college studies. For JUMP, five teachers provided unqualified positive comments, four made qualified positive comments, and one commented negatively. Positive aspects included raised expectations for pupils, good targeting of weaker pupils, and pupils' attitudes to having a different set of materials to the rest of the school. The qualification comments typically related to the need to adapt materials to an Irish context, or to spend more than a single year using JUMP in order to see real benefits. The sole negative comment related to a perceived lack of repetition of number facts and insufficient problem-solving activities.

"There is a lot of revision so it takes a long time to cover the course, but the kids liked having the book that 'nobody else had'. The second book was billed as 'the hard book' so they felt they were progressing when they did it!"

"It's an interesting programme and it'd be great to see it directly linked with the curriculum."

Suggestions for improving the evaluation

In the second set of interviews, teachers were asked to suggest changes in how the evaluation was organised. All had suggestions for improving the organisation of the evaluation. The majority related to lack of initial training on the programme, the limited amount of ongoing support and guidance provided, and organisational difficulties. Face-to-face contact with trainers and informal contact with other teachers were cited as important to counter teacher isolation.

Nine JUMP (and eight IMPACT) teachers said that all participating teachers needed to attend the initial CPD session, while 10 JUMP and nine IMPACT teachers felt that more ongoing support and contact was needed. Specific to programme, three JUMP teachers wanted more advice on pacing and how to cover the large volume of material, while three IMPACT teachers wanted less theoretical, and more practical, advice.

"*It was too long at the beginning without any support – that's a huge negative for a new programme [JUMP]. Pacing was an issue. We needed to liaise with other teachers much sooner.*"

"You need a liaison person for the programme – an expert on JUMP. John Mighton is too far away and not familiar with the Irish system."

"We need more contact and support in general and more practical, hands-on help at the start."

As might be expected, a few teachers believed that the very short notice for the initial training sessions was an indicator that the entire evaluation was planned with little notice.

"It was too last-minute. They didn't have it planned out."

Others complained that they were given insufficient detail about what participation in the evaluation involved, either in terms of general supports to be provided[7] or in terms of an indicative timetable of how to progress through the programmes.

"I'd like an outline schedule at the start of the year, indicating what is going to happen, when."

"They need a suggested overview plan for the year – to use as a guide for when to do specific bits of the programme. Say month-by-month or half-term by half-term."

Summary

In many ways, teacher descriptions of typical lesson activities did not vary too much by programme. Materials such as manipulatives, games, and real-life materials were ubiquitous. Textbooks were the basis for almost all lessons at the start of the year, and for a large proportion of lessons in May, although JUMP teachers were more likely to use workbooks in most or all lessons. Some of the lesson strategies that might be considered key JUMP principles – breaking ideas down into small steps, regular individual work, checking pupil understanding, pupil memorisation of rules and procedures – did not differ noticeably by programme, or differences at the end of the year were no more pronounced than pre-existing differences.

For most teachers in each programme, their self-rated preparedness for teaching PSMC strands, and the extent to which they felt confident about various aspects of mathematics instruction, increased over the year. This may be due to programme effects, to a greater concentration of effort on mathematics during the year, or simply because the initial ratings were depressed because they were given shortly after completing the MKTQ-S. However, a somewhat unexpected finding was the drop in the number of JUMP teachers expressing confidence in their ability to adapt their teaching to engage pupils' interest (dropping from 38% to only 13% indicating they were *very confident)*.

[7] In fact, the initial training day began with an explanation of the overall evaluation structure, and a description of the supports to be provided to teachers. However, the sizeable number who did not attend the initial day did not receive this information, and it may not have been remembered by all who did attend.

Nonetheless, most teachers were positive about many aspects of their assigned programme. The quality and depth of JUMP materials were cited as a positive, as was the hands-on approach, and the emphasis on building pupil confidence. Criticisms of JUMP typically centred on the need for adaptation to an Irish context, and organisational problems in the pilot study (especially in relation to CPD). Nearly half of teachers in each programme did not attend the initial CPD day that introduced teachers to the main features of the programme they were to implement during the year. Of these non-attending teachers, five did not take part in the November webinar. Thus, three IMPACT and two JUMP teachers spent half of the academic year without any training for their assigned programme. In sum, initial delays regarding elements of project funding contributed to very short notice for CPD, which contributed to poor uptake of CPD, and probably to poorer programme implementation within the classroom.

In addition, dissatisfaction was expressed with the amount of CPD provided and with webinar formats generally. Unlike issues around poor attendance, the amount and nature of CPD provided was unrelated to administrative problems. The training design (an initial training day supplemented by webinars) was based on advice from JUMP staff, but was perceived as insufficient by many Irish teachers. Perhaps due to initial lack of clarity about programme aims and methodologies, only a minority of JUMP teachers used the Confidence Building Unit as intended. About half of JUMP teachers used JUMP in conjunction with other programmes. In particular, a significant minority of teachers used other materials when teaching much of the Measures and Data strands.

Chapter 7
Mathematics achievement

This chapter provides information on pupil mathematics achievement at the start of the evaluation in September 2013 and in May-June 2014.[1] The main function of the chapter is to examine the extent and nature of change in mathematics achievement (if any) between September and May within each programme, overall, and for subsets of pupils.

The chapter is divided into seven sections, the first of which deals with overall mean scores of pupils, by programme. Section two analyses results by content and process. Section three examines achievement results by gender, while the fourth section examines how high- and low-achieving pupils performed in each programme. Section five examines pupil achievement by assigned observation type (i.e., recorded or live observations). The sixth section examines how teacher and teaching characteristics (i.e., teachers' mathematical knowledge for teaching, mathematical quality of instruction, and programme adherence) relate to pupil achievement, and if teachers' mathematical knowledge for teaching changed during the evaluation. Finally, the relationship between pupil achievement and attitudes to mathematics is examined. The analyses reported in the chapter are unadjusted for multiple comparisons, but where adjusted *p* values alter significance, this is noted at the end of the chapter.

Overall score on the DPMT

In September 2013, 546 pupils completed the DPMT, Level 2 (autumn norms). In May 2014, 536 pupils completed the DPMT, Level 3 (spring norms). In total, a "core group" of 509 pupils (271 in JUMP and 238 in IMPACT) completed the test on both occasions. Unless otherwise stated, all analyses that follow refer to the core group, or to a relevant subgroup within it. This allows a like-with-like comparison, as the same group of pupils are being compared on each occasion. However, readers should note that the mean score for pupils in the core group was slightly higher (on both occasions) than the mean score for all pupils sitting the test. This is an expected finding, as pupils absent for testing generally tend to have slightly lower achievement than those present for testing (see Cosgrove, 2005, for a discussion of the issue).

In September 2013, pupils in the core JUMP and IMPACT groups achieved mean standardised scale scores of 102, equivalent to the 55th percentile on the DPMT (Table 7.1). The similarity between the two groups was expected, given that standardised test scores from the end of the 2012/13 year had been used to assign schools to the two programmes in a way that maximised similarity in pupil achievement. Mean scores from the school-administered tests at Second class had averaged 109 in both groups, suggesting a drop in scores over the summer. However, this may be an artefact finding, due to a combination of different tests being used at the end of the previous year, pupil nervousness in September due to being tested by an external test administrator, the absence of a "halo effect"[2] when test questions were scored by ERC staff, and September testing taking place unusually early in the school year. This early timing

[1] As almost all pupils tested on the second occasion were tested in May, with only a small number tested during the first days of June, the second test period will hereafter be referred to as the May testing period.

[2] The halo effect refers (in this instance) to a tendency for teachers to give their pupils the benefit of the doubt on test items with ambiguous responses. Teachers may draw on what they already know about pupil knowledge of a topic and score accordingly, rather than scoring purely based on the response provided in the test.

avoided programme effects, but increased summer learning loss effects, which can be pronounced for mathematics (Cooper, Nye, Charlton, Lindsay & Greathouse, 1996). Irrespective of the cause of the change from May to September 2013, the most important aspect – that the two groups were well matched on average achievement – remained unchanged.

Table 7.1: Mean achievement scale scores (and standard deviations) for the core group of pupils, by programme and test time

	JUMP (N=271)	IMPACT (N=238)
Sept. 2013	102.0 (13.6)	102.2 (15.4)
May 2014	108.9 (14.4)	107.5 (17.8)
Change	+6.9 (8.8)	+5.3 (9.8)

The September 2013 mean scores of pupils in both programmes were slightly, but significantly, higher than the DPMT mean score of 100 (based on a national standardisation sample from 2006) ([JUMP: $t(270)=2.36$, $p=.019$]; [IMPACT: $t(237)=2.15$, $p=.033$]). The slight difference could be attributed to the self-selected nature of schools involved, to the fact that pupils absent on either test occasion were excluded from the mean, to improved mathematics achievement nationally since 2006, or (probably) to a combination of these factors.

In May 2014, mean scores for the core group of pupils increased in both programmes, to 109 for JUMP pupils (equivalent to the 73rd percentile), and to 107 for IMPACT pupils (equivalent to the 68th percentile). For both groups, the end-of-year results were again significantly higher than the DPMT mean of 100 obtained by the standardisation sample in 2005 ([JUMP: $t(270)=10.13$, $p<.001$]; [IMPACT: $t(237)=6.46$, $p<.001$)]).

Within each group, increases in pupil achievement over the course of the evaluation were significant ([JUMP: $t(270)=-12.94$, $p<.001$], [IMPACT: $t(237)=-8.33$, $p<.001$]). The standard deviation for the DPMT is 15 scale score points, meaning that the JUMP group improved by just under half a standard deviation, and the IMPACT group by roughly one third of a standard deviation. However, the difference in achievement scores *between* the JUMP and IMPACT groups was not significant at either test time ([September: $t(507) =-0.15$, $p=.880$], [May: $t(507)=0.98$, $p=.328$]). To correct for the fact that the samples were clustered by school, a Minimum Detectable Effect Size (MDES) was calculated (Hutchinson and Styles, 2010; Bloom, 1995) – i.e., the smallest gap that would produce a statistically significant difference between JUMP and IMPACT in the end-of-year tests.[3] The MDES was calculated at 6.1 scale score points. Since the JUMP and IMPACT groups differed by only 1.4 points in May, the difference between them was not statistically significant.

In sum, the achievement scores of JUMP pupils improved by an average of almost seven scale score points between the start and end of the evaluation, while those of IMPACT pupils improved by an average of just over five points. Pupils in each programme scored significantly higher than the DPMT mean of 100 on both occasions, and within each programme, the increase over the course of the evaluation was significant. However, pupils in the two programmes were not significantly different *from each other* in September (as was intended) or in May (when differences might have been expected if one programme was markedly more effective than the other).

[3] This calculation of the MDES assumes a two-tailed test, $p<.05$, power=80% (Bloom, 1995).

Performance on strands and content areas

The DPMT items are subdivided into four content groups that reflect the strands of the PSMC: Number and Algebra, Shape and Space, Measures, and Data. All items are also classified as requiring one of five processes: Understanding and Recalling, Implementing, Reasoning, Integrating and Connecting, and Applying and Problem-Solving. Standardised scale scores are not available at the content and process level, meaning we can only report *percent correct scores*. In practical terms, this means that content and process results should not be compared with each other or across Levels of the DPMT. For example, answering 66% of Data items and 55% of Measures items correctly at Level 2 is not equivalent to doing the same at Level 3, nor does it imply the pupil is better on Data than Measures. The Measures questions answered by the pupils may have been more difficult than the Data questions – something taken into account when reporting standardised scale scores, but not by a percent correct score. Therefore, content and process results are compared between JUMP and IMPACT groups in September and in May, but are not compared between the two times.

Content

At the start of the evaluation, the JUMP and IMPACT groups were very closely matched on the percent of items answered correctly, by content area (Table 7.2). The difference was largest for Shape and Space (an advantage of 2.4% in favour of IMPACT pupils) and smallest for Data (0.5% in favour of IMPACT).

Table 7.2: Pupils' mean percent correct scores by content strand and programme, **September 2013**

	JUMP	IMPACT	Gap
Number and Algebra	64.7	63.1	1.6%
Shape and Space	67.1	69.5	2.4%
Measures	45.5	47.7	2.2%
Data	50.1	50.6	0.5%

At the end of the year, the mean percentage of items answered correctly for three of four content areas was again very similar across the two programmes. The exception was Data, where JUMP pupils achieved a mean percent correct score 4% higher than that obtained by IMPACT pupils (Table 7.3). The difference on Data is nonetheless relatively small, and is almost entirely attributable to differences in performance on three items assessing pupil ability to interpret a pictogram. On these items, the percentage of JUMP pupils answering correctly was considerably higher than the percentage of IMPACT pupils who did so (differences of 22%, 13%, and 19%). Of course, as the Data strand was not covered in the available IMPACT manuals, it was probably taught to most IMPACT pupils using commercially available Irish textbooks.

Table 7.3: Pupils' mean percent correct scores by content strand and programme, **May 2014**

	JUMP	IMPACT	Gap
Number and Algebra	62.0	59.8	2.3%
Shape and Space	67.6	66.9	0.7%
Measures	56.9	55.2	1.7%
Data	66.9	62.9	4.0%

Process

At the outset, pupils' percent correct scores on all processes were very similar across the JUMP and IMPACT groups (Table 7.4). The largest difference was on Applying and Problem-Solving (a 2% gap in favour of IMPACT), while differences on Understanding and Recalling, Reasoning, and Integrating and Connecting were all below 1%.

Table 7.4: Pupils' mean percent correct scores by process and programme, September 2013

	JUMP	IMPACT	Gap
Understanding and Recalling	66.3	66.9	0.6
Implementing	66.4	64.6	1.7
Reasoning	62.6	61.9	0.7
Integrating and Connecting	58.7	59.5	0.8
Applying and Problem Solving	45.3	47.3	2.0

At the end of the year, the groups were still closely matched across four of the processes. However, on Integrating and Connecting, JUMP pupils achieved a mean percent correct score almost 6% higher than that of IMPACT pupils (Table 7.5). Again, this can be partially attributed to large differences between JUMP and IMPACT pupils on the three items relating to pictograms (Integrating and Connecting was required for the three questions).

Table 7.5: Pupils' mean percent correct scores by process and programme, May 2014

	JUMP	IMPACT	Gap
Understanding and Recalling	72.8	71.2	1.6
Implementing	66.5	64.1	2.4
Reasoning	64.6	63.8	0.8
Integrating and Connecting	58.0	52.2	5.8
Applying and Problem Solving	53.5	51.5	2.0

Achievement by gender

As with the overall sample of pupils taking part in the evaluation, there were slightly fewer girls than boys in the core group tested (222 girls and 287 boys). Girls were outnumbered by boys in both JUMP (113 girls and 158 boys) and IMPACT (109 girls and 129 boys). Thus, there were fairly similar gender representations in the two groups, as girls composed 42% of pupils in JUMP and 46% in IMPACT.

In September, girls in both programmes achieved mean scores of 101, while boys in both programmes achieved mean scores of 103, a non-significant gender gap ($t(507)=-1.59$, $p=.112$). In May, boys again obtained slightly, but not significantly, higher mean scores than girls ($t(507)=-0.99$, $p=.322$). Girls and boys in JUMP achieved mean scores of 108 and 109, respectively, while girls and boys in IMPACT achieved scores of 107 and 108, respectively (Table 7.6). Thus, girls and boys in JUMP had an average increase of about seven points on the DPMT, while girls and boys in IMPACT both improved by between five and six points. These increases were significant for girls and boys in both programmes ([JUMP girls: $t(112)=-8.61$, $p<.001$], [JUMP boys : $t(157)=-9.64$, $p<.001$], [IMPACT girls: $t(108)=-6.15$, $p<.001$], [IMPACT boys: $t(128)=-5.67$, $p<.001$]).

Within each programme, gender differences were not significant in September ([JUMP: $t(269)=-1.12$, $p=.261$], [IMPACT: $t(236)=-1.13$, $p=.259$]) or May ([JUMP: $t(269)=-0.61$, $p=.542$], [IMPACT: $t(236)=-0.73$, $p=.468$]). Gender differences *across* programmes were also non-significant in September ([girls: $t(220)=-0.03$, $p=.973$], [boys: $t(285)=-0.26$, $p=.792$]) and May ([girls: $t(208)=0.76$, $p=.449$], [boys: $t(245)=0.58$, $p=.563$]). In sum, boys and girls in both programmes improved their mean achievement scores by a statistically significant amount, and neither programme showed differential effectiveness by gender.

Table 7.6: Mean achievement scores (and standard deviations) by gender, programme and time, and change in score

	JUMP (N=271)		IMPACT (N=238)	
	Girls	Boys	Girls	Boys
Sept. 2013	100.8 (13.7)	102.7 (13.6)	100.9 (15.9)	103.2 (15.0)
May 2014	108.2 (14.8)	109.3 (14.1)	106.5 (18.3)	108.2 (17.3)
Change	+7.4 (9.1)	+6.6 (8.6)	+5.6 (9.5)	+5.0 (10.1)

Results of initially low- and high-achieving pupils

This section analyses the results of two subsets of pupils:

- Low-achieving pupils, operationally defined as those whose DPMT score in September 2013 was more than one standard deviation (15 scale score points) *below* the national mean.
- High-achieving pupils, operationally defined as those whose DPMT score in September 2013 was more than one standard deviation *above* the national mean.

In the initial administration of the DPMT, 53 pupils from the core group achieved scores more than one standard deviation lower than the national mean, while 107 pupils achieved scores more than one standard deviation above the mean (Table 7.7). The 22 low-achieving JUMP pupils had a mean score of 80, and the 31 IMPACT pupils had a significantly lower mean score of 77 ($t(49)=2.64$, $p=.011$). The 56 high-achieving JUMP pupils obtained a mean of 122, very similar to that of 123 obtained by the 51 high-achieving IMPACT pupils. Thus, both groups were closely matched on numbers and mean scores of high-achieving pupils, but IMPACT had slightly more low achievers, with a slightly lower mean score (80 versus 77).

Looking at how the same pupils performed in May 2014, low-achieving pupils in both programmes improved by about five score points (one third of a standard deviation). These increases were significant in both programmes ([JUMP: $t(21)=-2.81$, $p=.011$], [IMPACT: $t(30)=-3.87$, $p=.001$]). The difference across programmes between scores of low-achieving pupils was no longer statistically significant at the end of the year ($t(51)=1.20$, $p=.236$).

Over the course of the evaluation, the scores of high-achievers also improved (significantly), by two points in JUMP and three points in IMPACT ([JUMP: $t(55)=-2.50$, $p=.015$], [IMPACT: $t(50)=-2.56$, $p=.013$]). However, the results of high-achieving JUMP pupils did not differ significantly from those of high-achieving IMPACT pupils in September ($t(105)=1.10$, $p=.274$) or May ($t(86)=1.38$, $p=.173$).

The number of cases is very small when achievement and programme are further split by gender (ranging from only 10 girls in the JUMP low-achieving group to 35 boys in the JUMP high-achieving group). Given the small numbers involved, tests of statistical significance were not deemed appropriate. However, the mean scores can still be described, once interpreted with caution (Table 7.7). Examining low-achieving pupils by gender, girls in both JUMP and IMPACT achieved lower mean scores than boys on both occasions. The JUMP gender gap remained almost identical (2.9 points at the start of the year and 2.8 points at the end), while the IMPACT gender gap increased slightly (from 2.3 points to 3.9 points). Examining high-achieving pupils by gender, girls in both programmes began in September with slightly lower mean scores than boys in September, but ended with slightly higher scores than boys in May. Among high achievers, the scores of IMPACT girls increased the most (4.1 points).

Table 7.7: Mean achievement scores (and standard deviations) for pupils whose baseline scores were more than one standard deviation below/above the mean, overall and by gender, programme and test time

		JUMP			IMPACT		
		Sept '13	May '14	Diff	Sept '13	May '14	Diff
1 SD below	Total (N=53)	80.0 (3.1)	85.5 (9.8)	+5.5 (9.3)	76.8 (5.5)	82.1 (10.3)	+5.3 (7.7)
	Girls (N=25)	78.4 (3.7)	84.0 (9.4)	+5.6 (8.4)	75.6 (5.8)	80.1 (9.3)	+4.5 (6.1)
	Boys (N=28)	81.2 (1.7)	86.7 (10.3)	+5.5 (10.3)	77.9 (5.2)	84.0 (11.2)	+6.1 (9.0)
1 SD above	Total (N=107)	122.1 (6.1)	124.2 (6.5)	+2.1 (6.4)	123.4 (6.1)	126.4 (9.6)	+3.0 (8.5)
	Girls (N=43)	121.9 (5.0)	124.4 (7.3)	+2.6 (7.1)	122.9 (5.6)	127.0 (9.3)	+4.1 (8.4)
	Boys (N=64)	122.2 (6.7)	124.1 (6.0)	+1.9 (6.0)	123.7 (6.6)	125.9 (10.0)	+2.2 (8.6)

Observation type and pupil achievement

As described in Chapter 4, mathematics lessons were observed "live" in 11 classes (on two occasions) and recorded in 15 classes (also on two occasions). Many teachers indicated that they were particularly nervous about being recorded while teaching, while some also requested copies of the recorded lessons. On the basis that either the act of being recorded or the effects of watching a recorded lesson might affect teaching behaviour (e.g., improved teaching due to reflecting on observed teaching behaviour, extra effort due to being recorded, poorer quality teaching due to the stress of being recorded), pupil results were analysed by observation type.

JUMP pupils in recorded classes had a higher mean increase in scores than JUMP pupils in non-recorded classes (eight score points as opposed to six score points). For IMPACT pupils, the mean baseline score was higher among pupils in non-recorded classes, but there was almost no difference in mean increases between those who were recorded and those who were not (Table 7.8). Significant increases in mean scores were evident in both recorded and non-recorded groups in JUMP ([recorded: $t(121)=-11.07$, $p<.001$], [non-recorded: $t(148)=-7.75$, $p<.001$]) and IMPACT ([recorded: $t(134)=-5.60$, $p<.001$], [non-recorded: $t(102)=-6.38$, $p<.001$]). However, the difference in the size of the increase *between* recorded and non-recorded groups was negligible for IMPACT, and not significant for JUMP on either occasion ([September: $t(269)=-.37$, $p=.713$], [May: $t(269)=1.21$, $p=.227$]). In other words, observation type did not appear to have a significant effect on pupils' achievement results.

Table 7.8: Mean achievement scores and standard deviations, by observation type, programme, and test time

	Recorded observation		Non-recorded observation	
	JUMP (N=122)	IMPACT (N=135)	JUMP (N=149)	IMPACT (N=103)
Sept 2013	101.6 (14.6)	100.0 (15.5)	102.2 (12.9)	105.0 (14.8)
May 2014	110.0 (14.5)	105.0 (18.2)	107.9 (14.3)	110.7 (16.8)
Change	+8.4	+5.0	+5.7	+5.7

Teacher variables and pupil achievement

In this section, pupils' achievement gains are analysed vis-à-vis four teacher characteristics:

- Mathematical knowledge for teaching (MKT).
- Mathematical quality of instruction (MQI).
- CPD uptake.
- Programme adherence.

Mathematical knowledge for teaching

The shortened Mathematical Knowledge for Teaching Questionnaire (MKTQ-S) was completed by 31 teachers in September 2013, and by 26 teachers in May 2014. Data from the MKTQ-S were used for two purposes, bulleted below and described in the next sections.

- to assess whether teachers' mathematical knowledge for teaching changed during the year, using the results of a core group of teachers who sat the test on both occasions.
- to assess whether pupils' achievement was related to their teacher score on the MKTQ-S, using baseline teacher results where available.

Overall scores on the MKTQ-S

The core group that completed the MKTQ-S on both occasions comprised 23 teachers: 14 from JUMP (representing 11 classes) and nine from IMPACT (representing nine classes). The small numbers reflect a slight drop in response rate at the end of the evaluation, and the difficulty in "matching" data for teachers who were job-sharing, on leave, or working as learning support teachers in a number of classes. Scores are reported as the number of questions answered correctly, relative to the Irish norm group of teachers (Delaney, 2012).

In September 2013, the MKTQ-S scores of core group teachers in JUMP and IMPACT differed by less than one question answered correctly (Table 7.9). Both figures are slightly higher than the mean score achieved on MKTQ-S questions by a previous sample of 500 Irish teachers, who were administered the whole MKTQ (S. Delaney, personal communication, August 2013). This may reflect the self-selection of teachers interested in mathematics, as well the shorter length of the MKTQ-S.[4]

In May 2014, teachers in both programmes answered an average of roughly four and a half more questions correctly than did the Irish norm group. This is equivalent to a 10% advantage on the MKTQ-S for teachers in the evaluation, relative to the norm group. However, there was considerable individual variation in both groups. The number of questions answered correctly by two teachers (one in each programme) decreased by five, while questions answered correctly increased by nine for two JUMP teachers and by 17 for one IMPACT teacher. However, in most cases, the difference in start- and end-of-year scores ranged between a decrease of one and an increase of four correct answers.

Table 7.9: Difference in the number of MKTQ-S questions answered correctly by teachers in the core group (by programme and time) versus Delaney's Irish norm group

JUMP (N=14)		IMPACT (N=9)	
Time 1	Time 2	Time 1	Time 2
+1.4	+4.4	+2.1	+4.3

[4] Although the MKTQ-S is a halved form of the MKTQ, it usually took at least 40 minutes to complete. As such, it is possible that fatigue may depress scores for the full version.

Teachers' MKTQ-S and pupil achievement

Teacher MKTQ-S scores were applied to classes, using teachers' results from September where possible, to provide a baseline measure of their mathematical knowledge for teaching. For three teachers for whom September data were not available, May results were used. The mean score for JUMP classes was two correct answers fewer than for IMPACT classes.

When mean class MKTQ-S scores were correlated with gains in pupil achievement (Pearson, two-tailed), a small but statistically significant *negative* correlation was found for JUMP (r=-.17, N=271, p=.005) (Table 7.10). That is, JUMP classes with lower mean MKTQ-S scores tended to have slightly *higher* gains in pupil achievement over the course of the evaluation. However, this correlation represents a very small percentage of variance (2.9%). There was not a significant correlation for IMPACT. Correlations between mean class MKTQ-S scores and end-of-year pupil achievement scores were not significant for either programme.

Table 7.10: Correlation between baseline class MKTQ-S score and: (a) pupil difference score; (b) pupil achievement score in May 2014

	JUMP (N=271)	IMPACT (N=238)
r for pupil difference score and class MKTQ-S score	- 0.17**	0.03
r for pupil May 2014 DPMT score and class MKTQ-S score	0.04	0.10

** Significant at 0.01 level.

Mathematical quality of instruction (MQI)

Recorded observations were conducted at two points during the year for 15 classes, and were rated on the MQI by two observers on each occasion (see Chapter 4). As some teachers indicated that they were nervous and/or taught an atypical lesson for the first observations, ratings from the second observations were considered more reflective of teachers' general practice. Therefore, drawing on the two observers' overall MQI ratings from the second observation *only*, the mean rating was 1.9 for JUMP and 2.1 for IMPACT. As the MQI is a three-point scale (where 1 equals low quality and 3 equals high quality), both means might be considered to reflect medium quality levels.

Class MQI ratings were correlated (Pearson, two-tailed) with gains in pupil achievement and with pupils' May scores (Table 7.11). Neither correlation was significant for JUMP pupils, but both were significant for IMPACT pupils ([MQI and gains in pupil achievement: r=.25, N=135, p=.004, corresponding to 6% of variance], [MQI and May pupil achievement: r=.40, N=135, p<.001, corresponding to 16% of variance]). Thus, gains and overall achievement scores were higher for pupils taught by IMPACT teachers who were rated as displaying higher quality of instruction.

Table 7.11: Correlation between class MQI rating from second observations, and: (a) gains in pupil achievement; (b) pupil achievement score in May 2014

	JUMP (N=122)	IMPACT (N=135)
r for pupil difference score and MQI rating	-0.04	0.25**
r for pupil May 2014 DPMT score and MQI rating	0.12	0.40**

**Significant at 0.01 level.

CPD uptake

Linking pupil achievement and teacher uptake of CPD is somewhat problematic. There are qualitative differences between attending the initial training day, viewing it online at a later date, attending some or all of a webinar, or doing all of the preceding. Thus, CPD attendance cannot simply be summed and correlated with achievement, meaning teacher uptake of CPD falls into multiple (not directly comparable) categories. Given a small number of teachers, already divided by programme, most categories of CPD fall below the point where analyses of statistical significance are appropriate. Therefore, this section only examines teacher attendance at the initial training day. Attendance is defined as physical attendance, and does not include the small number who later watched recordings of part or all of the initial sessions.

Teachers from nine JUMP and nine IMPACT classes attended the initial training day.[5] JUMP pupils whose teachers attended initial training showed an increase of 7.7 points in mean scores from September 2013 to May 2014, while JUMP pupils whose teachers had not attended showed an increase of 5.0 points. The difference in gains between these groups was statistically significant ($t(269)=2.34$, $p=.020$). Amongst IMPACT pupils, the average increase in scores was 4.9 points for pupils whose teachers attended, and 5.7 points for those whose teachers did not. This difference was not statistically significant ($t(236)=-0.62$, $p=.539$).

Programme adherence

As with MQI ratings, programme adherence ratings from the second observations were thought to be more reliable than those from the first observations. No significant correlations were found between programme adherence and gains in pupil achievement (Table 7.12). Programme adherence was significantly positively correlated with IMPACT pupils' achievement scores in May ($r=.17$, N=225, $p=.009$).

Table 7.12: Correlation between class adherence rating from second observations, and: (a) pupil difference score; (b) pupil achievement score in May 2014

	JUMP (N=271)	IMPACT (N=225)
r for pupil difference score and adherence	-0.09	0.04
r for pupil May 2014 DPMT score and adherence	0.05	0.17**

**Significant at 0.01 level.

The sample of teachers is too small to warrant reporting the correlation values for adherence and other teacher variables, e.g., MKT baseline score, MKT difference score, and MQI rating (second observation). However, no strong associations were apparent between programme adherence and any of these variables.

Pupil attitudes and pupil achievement

This section examines the relationship between pupils' attitudes to mathematics (as reported in Chapter 5) and their mathematical achievement. Findings are grouped as relating to pupils':

- *general attitudes* to mathematics.
- *confidence* in their own mathematical competence.
- reports of their *teachers' behaviour* in mathematics class.
- *habits* when learning and practising mathematics.

[5] In a tenth IMPACT class, the teacher who attended initial training subsequently went on leave and was replaced by a new class teacher, who had not attended initial training.

General attitudes to mathematics

At the start and end of the year, pupils were asked to agree or disagree (a little or a lot) with the statements that they *liked mathematics* and that they *wished they didn't study mathematics.*

In September, pupils' liking of mathematics was significantly positively correlated with achievement in JUMP (r_s=.17, N=255, p=.007), but not IMPACT (r_s=.12, N=221, p=.080). Similarly, there was a significant negative correlation between pupil achievement and wishing they didn't study mathematics for JUMP (r_s=-.21, N=266, p<.001), but not IMPACT (r_s=-.02, N=227, p=.714). However, in May, this pattern was reversed. For IMPACT, achievement was significantly positively correlated with liking mathematics (r_s=.16, N=231, p=.013), but this was not the case for JUMP (r_s=.010, N=268, p=.095). Achievement in both groups was significantly negatively correlated with not wanting to study maths ([JUMP: r_s=-.15, N=270, p=.02], [IMPACT: r_s=.23, N=237, p<.001]) (Table 7.13). Thus, during the evaluation, the association between positive attitudes to mathematics and achieving high scores decreased in JUMP, but increased in IMPACT.

Table 7.13: Correlation between pupils' mean achievement scores and two variables indicating general attitude to mathematics, by programme and test time

	JUMP		IMPACT	
	Sep. '13	May '14	Sep. '13	May '14
r_s for achievement score and *I like maths*	0.17**	0.01	0.12	0.16
r_s for achievement score and *I wish I didn't study maths*	-0.21**	-0.15*	-0.02	-0.23**

*Significant at 0.05 level **Significant at 0.01 level

The decreased association between attitude and achievement in JUMP is explained in part by the fact that, in May, JUMP pupils who "disagreed a little" that they liked maths had generally high achievement scores (mean scaled score 112, range 94-133, N=29), as did those who "agreed a little" that they didn't want to study maths (mean score 110, range 84-136, N=75) (Table 7.14). These groups of pupils also had high mean increases in achievement, averaging to nine score points in both cases. Therefore, some JUMP pupils whose achievement scores *and* gains were high at the end of the evaluation had more negative attitudes to mathematics than might have been expected.

Table 7.14: Mean maths achievement scores for pupils who agreed to various extents with two statements concerning attitude to mathematics, by programme and test time

		JUMP		IMPACT	
		Sep. '13	May '14	Sep. '13	May '14
"I like maths"	Agree a lot	104.6	109.9	104.8	109.7
	Agree a little	102.8	108.9	103.7	109.6
	Disagree a little	97.4	111.7	101.6	105.5
	Disagree a lot	99.4	103.8	99.4	100.8
"I wish I didn't study maths"	Agree a lot	98.3	102.9	102.8	99.8
	Agree a little	100.4	110.1	101.3	105.3
	Disagree a little	103.5	109.8	104.2	110.6
	Disagree a lot	105.5	110.4	102.8	110.7

Confidence in mathematical ability

At the start and end of the year, pupils were asked to agree or disagree (a lot or a little) with statements relating to mathematical confidence and anxiety: including that that they were good at maths, that they worried about not being able to answer questions in maths class, and that they thought everyone could be good at maths.

For pupils in both programmes, on both occasions, positive correlations between belief in being good at maths and achievement score were significant, although slightly larger in September ([JUMP: r_s=.29, N=263, p<.001], [IMPACT: r_s=.33, N=231, p<.001]) than in May ([JUMP: r_s=.22, N=271, p<.001] [IMPACT: r_s=.27, N=236, p<.001]) (Table 7.15).

There was also a significant negative correlation between achievement scores and the extent to which pupils agreed with the statement *I worry I won't be able to answer questions in maths class*, in both programmes and on both occasions. The size of the correlation decreased slightly for JUMP, but increased for IMPACT, from September ([JUMP: r_s=-.31, N=264, p<.001], [IMPACT: r_s=-.17, N=229, p=.01]) to May ([JUMP: r_s=-.21, N=266, p<.001], [IMPACT: r_s=-.35, N=237, p<.001]). In other words, for IMPACT, the association between low achievement and pupil anxiety levels about being asked questions increased over the course of the year.

Table 7.15: Correlation between pupils' mean achievement scores and two variables indicating confidence in their own mathematical competence, by programme and test time

	JUMP		IMPACT	
	Sep. '13	May '14	Sep. '13	May '14
r_s for DPMT score and pupils agreeing *I am good at maths*	0.29**	0.22**	.33**	0.27**
r_s for DPMT score and *I worry I won't be able to answer questions in maths class*	-0.31**	-0.21**	-0.17*	-0.35**

*Significant at 0.05 level **Significant at 0.01 level

The belief that everyone can be good at maths was not significantly correlated with achievement for pupils in either programme, at either test time. However, among JUMP pupils, there was a significant positive correlation at the start of the year between how much they believed *they* were good at mathematics and how much they believed *everyone* could be good at maths (r_s=.23, N=251, p<.001). At the end of the year, the correlation was not significant (r_s=.10, N=267, p=.107) (Table 7.16). For IMPACT pupils, there was no significant correlation, on either occasion, between belief in their own mathematical ability and belief in the potential of everyone to be good at mathematics.

Table 7.16: Correlation between pupils' belief that *everyone could be good at maths* and (a) achievement scores; (b) belief that *they are good at maths*, by programme and test time

	JUMP		IMPACT	
	Sep. '13	May '14	Sep. '13	May '14
r_s for achievement score and pupils agreeing that *everyone can be good at maths*	-0.04	0.02	-0.04	-0.08
r_s for pupils believing *I am good at maths* and that *everyone can be good at maths*	0.23**	0.08	0.10	0.04

**Significant at 0.01 level

Teachers' behaviour in mathematics class

Pupils were asked to agree or disagree (a lot or a little) with statements about what their teacher did in their mathematics lessons. Those that were related to achievement are shown in Table 7.17. In September, pupils' agreement that their teacher *always asked if they understood* was positively correlated with achievement for both programmes ([JUMP: r_s=.13, N=266, p=.043], [IMPACT: r_s=.18, N=227, p=.005]). However, this correlation was not significant for either programme in May.

JUMP pupils' agreement that their teacher *got them to practice lots of examples* was *negatively* correlated with achievement at both test times, with correlation size increasing on the second occasion ([September: r_s=-.14, N=266, p=.020], [May: r_s=-.24, N=267, p<.001]). The correlation was not significant for IMPACT pupils at either test time.

At the outset, pupils' beliefs that their teacher *gave them fun things to do* and *let them play games* were not significantly correlated with achievement for either programme. However, at the end of the year, agreement that the teacher gave them fun things to do was *negatively* correlated with achievement in both groups ([JUMP: r_s=-.17, N=270, p=.006], [IMPACT: r_s=-.13, N=236, p=.049]). For JUMP pupils, agreement that the teacher let them play games was also negatively correlated with achievement (r_s=-.20, N=271, p=.001), though this correlation was not significant in IMPACT.

At the end of the year, there was a small negative correlation between JUMP pupils' achievement and their agreement that their teacher *often praised them* (r_s=-0.14, N=266, p=.025). This correlation was not significant for IMPACT.

Table 7.17: Correlation between pupils' achievement scores and their agreement with various statements about their teachers' behaviour in mathematics class

	JUMP		IMPACT	
	Sep. '13	May '14	Sep. '13	May '14
"My teacher always asks do we understand stuff."	0.13*	-0.06	0.18**	0.10
"My teacher often praises me."	0.04	-0.14*	0.07	-0.07
"My teacher gets me to practice lots of examples."	-0.14*	-0.24**	0.07	-0.02
"My teacher gives us fun things to do."	0.50	-0.17**	0.13	-0.13*
"My teacher lets us play games."	0.01	-0.20**	0.07	0.07

*Significant at 0.05 level **Significant at 0.01 level

Habits when learning mathematics

Pupils were asked to indicate how frequently they used various learning strategies for mathematics (*every class*, *most classes*, *some classes*, or *hardly ever*). Table 7.18 shows those significantly correlated with achievement.

In September, frequency of repeating examples was not significantly correlated with achievement for either group, but in May there was a significant *negative* correlation for both JUMP (r_s=-.20, N=265, p=.001) and IMPACT (r_s=-0.18, N=234, p=.006). That is, at the end of the year, pupils who reported frequently repeating examples were likely to have lower achievement scores than those who did not.

At the start of the year, there was a significant negative correlation between achievement and how often pupils worked with their classmates to solve a problem for JUMP (r_s=-0.18, N=258, p=.005), but not IMPACT. At the end of the year, however, a negative correlation was significant for IMPACT (r_s=-.16, N=229, p=.019), but not JUMP. For the item asking how often pupils worked out a sum in their heads, a significant negative correlation with achievement was observed for IMPACT pupils in May (r_s=-.15, N=233, p=.020).

Table 7.18: Correlation between pupils' achievement scores and the frequency with which pupils reported engaging in various learning strategies for mathematics

	JUMP		IMPACT	
	Time 1	Time 2	Time 1	Time 2
"When we do new things, I learn as much as I can by heart."	0.16**	0.07	0.04	-0.12
"I go through examples again and again to help me remember them."	-0.06	-0.20**	0.01	-0.18**
"I work with my classmates to solve a problem."	-0.18**	-0.11	-0.07	-0.16*
"I work out a sum in my head."	0.04	-0.05	0.03	-0.15*

*Significant at 0.05 level **Significant at 0.01 level

Summary

Pupils' achievement improved significantly in *both* programmes over the course of the evaluation. On average, JUMP pupils improved slightly more than IMPACT pupils (seven scale score points as opposed to five), but this difference was not statistically significant.

When results were broken down by strand and process, there were few programme-based differences. However, in May, JUMP pupils outperformed IMPACT pupils on the Data strand and on the process of Integrating and Connecting. These differences are largely attributable to pupils' performance on a few linked items relating to pictogram interpretation.

Gender differences were small, with boys slightly, but not significantly, outperforming girls across the whole sample and within the two programmes on both occasions. Achievement gains were significant for both girls and boys in the two groups.

In both programmes, pupils who achieved low scores in the September tests gained about five points by the May tests, on average. Pupils who achieved high scores at the start of the evaluation gained about two points in JUMP and three in IMPACT.

Teachers' scores on the MKTQ-S improved slightly in both JUMP and IMPACT during the evaluation. However, teachers' (initial) Mathematical Knowledge for Teaching scores were not related to pupils' end-of-year achievement scores. JUMP teachers' MKT scores were negatively correlated with their pupils' *gains* in achievement, indicating that teachers with lower MKT scores saw bigger improvements in their pupils' performance over the year than did teachers with higher MKT scores.

Mathematical Quality of Instruction (as rated by subject matter experts) was unrelated to pupil achievement in JUMP. In IMPACT, however, there was a significant positive correlation between MQI and both pupils' end-of-year scores and gains in achievement. JUMP pupils whose teachers had attended the initial training day showed slightly, but significantly higher gains those whose teachers did not attend. No attendance effect was apparent for IMPACT. In contrast, teachers' adherence to programme ethos was not linked to achievement in JUMP, but was positively correlated with pupils' end-of-year scores in IMPACT.

At the start of the year, pupil achievement was significantly correlated with having a positive general attitude towards mathematics in JUMP, but not IMPACT. However, at the end of the year, this pattern was reversed, partly because high-achieving (and high-gaining) JUMP pupils now expressed slight dislike for mathematics. High confidence and low anxiety were correlated with achievement for pupils in both programmes, both at the start and end of the evaluation, although this pattern was stronger in IMPACT than JUMP. In particular, the relationship between low achievement and anxiety about being asked questions in class was more pronounced in IMPACT pupils at the end of the year.

Few teacher behaviours (as reported by pupils) were positively correlated with achievement. In May, achievement in both programmes was *negatively* correlated with teachers frequently "giving pupils fun things to do". In JUMP, it was also negatively correlated with teachers letting pupils play games, getting them to practice a lot of examples, and praising them.

In May, for both programmes, achievement was negatively correlated with pupils reporting frequent use of repeated examples in their own mathematics practice. In IMPACT, it was also negatively correlated with the frequency of pupils working with classmates, and working out sums in their heads.

Overall, the achievement results indicate significant improvements in both programmes, but do little to distinguish the programmes' effects *from one another*.

Note on significance levels

When more than two mean scores are simultaneously compared, there is an increased probability of what is called a "Type 1 error", or a false positive. That is, the more comparisons are made, the greater the likelihood that the groups will differ on at least one comparison.

The Dunn-Bonferroni procedure (see for instance Kirk, 1968) can be used to control for this possibility, by calculating an *adjusted* significance level. The significance level is divided by the number of related t-tests carried out, with the resulting figure divided by two for a two-tailed test. For instance, if the desired significance level for a single t-test is .05 (that is, a one in twenty chance of a Type 1 error), and eight related t-tests are carried out, the adjusted significance level is .003. This is obtained as follows: [$\frac{0.5}{8}$ = 0.006], then [$\frac{0.006}{2}$ = .003]. In this instance, each of the eight t-tests would only be statistically significant if the *p* value was less than .003, not the original .05.

The data reported throughout this chapter draw on uncorrected significance levels, in part because corrections for multiple comparisons, coupled with a relatively small sample, can lead to extremely conservative interpretation of significance. When the Dunn-Bonferroni adjustment is applied, four results move from significant to non-significant:

1. JUMP and IMPACT pupils' overall September scores are not significantly different from the population mean (adjusted significance level=.003).
2. The difference between low-achieving pupils in JUMP and IMPACT at the start of the year is not significant (adjusted significance level=.006).
3. The gains made by low-achieving JUMP pupils during the year are not significant (adjusted significance level=.006).
   However, the difference in *p* values for JUMP and IMPACT in this regard is due partly to the smaller number of JUMP than IMPACT pupils in the low-achieving category.
4. The gains made by high-achieving pupils in both programmes are not statistically significant (adjusted significance level=.006).

Chapter 8
Conclusions

The purpose of this chapter is to summarise and comment on the main outcomes of the JUMP evaluation conducted in 2013/14. The study used a variety of measures to evaluate JUMP, including analyses of JUMP materials and of how well JUMP matched the content objectives of the Irish primary school mathematics curriculum (PSMC). Classroom observations were conducted, pupils' and teachers' attitudes and behaviours were examined using questionnaires and interviews, and pupil mathematical achievement was assessed pre- and post-intervention using a standardised achievement test widely used in Irish primary schools (the Drumcondra Primary Mathematics Test).

Unlike most previous evaluations of JUMP, the evaluation included a comparison group (the IMPACT programme, developed by the PDST to help teachers implement the PSMC) and classes were assigned to each programme in a manner that maximised similarity between baseline pupil mathematics achievement. Previous studies of JUMP efficacy have tended to examine how JUMP works with a selection of pupils/classes/districts without a control group (e.g., Aduba, 2006, 2007, 2009; Maciejewski, 2012 [centring on college-age students]; Murray, 2013) or only to focus on pupils' mathematical attitudes or confidence (e.g., Biswas Mellamphy, 2004; Hughes, 2004). These studies have typically found positive effects arising from JUMP, but interpretation of such findings can be problematic. In particular, in Aduba's Lambeth study (where JUMP was mainly delivered as an add-on mathematics programme for low-achieving pupils), JUMP coincided with a number of other interventions, including intensive whole-school improvement programmes in many participating schools. This makes it difficult to disentangle the effects of JUMP from the effects of the school having extra support from a range of professionals (education advisers, teaching and learning consultants, outreach workers, partnership schools).

Solomon et al.'s (2011) as-yet-unpublished study used a random control trial model to assign classes to JUMP or to continue with the regular (Ontario) curriculum. While the randomised assignment to programme is preferable to school self-selection or being "volunteered" for the programme as part of a wider attempt to address school underperformance, it does not seem to address the fact that the control group knew they were the control group. Garforth's (2013) work is an exception, in that she compared classes who experienced JUMP alone, JUMP combined with other programmes, and normal curriculum materials. Her results indicated that JUMP was associated with improved performance only for Number, and poorer performance on computational skills. However, with only 106 pupils across the three conditions, it is difficult to draw many conclusions from such a small sample.

Therefore, as noted in Chapter 1, this evaluation included a control (or comparison) group, with efforts made to ensure they did not know they were the control group. Although IMPACT was not the main focus of the evaluation, significant parts of the preceding chapters have described what happened in IMPACT classes. This was to take into account two key issues that might affect attitudes and behaviour: the Hawthorne Effect (possible effects on behaviour from knowing you are part of an evaluation), and gradual changes that might occur naturally over the course of an academic year. Thus, any changes that manifested only in the JUMP group might therefore be more directly attributable to characteristics of the JUMP programme.

Unlike previous chapters, the focus here is largely on JUMP. The chapter is divided into three sections, the first of which outlines the five main questions the evaluation was designed to answer, and what answers were found. The second section examines what contributed to improved mathematics achievement, and the third section examines future possibilities for JUMP in Ireland, drawing on the outcomes of the present evaluation.

All of what follows is, however, necessarily limited by the relatively small scale of the study. The overall number of pupils is quite small, the numbers of schools and teachers are very small, and analyses by subgroups of pupils within each programme are therefore somewhat problematic. These considerations should be borne in mind when interpreting findings.

The research questions

There were five main research questions of interest in the evaluation. The first, and most important, was the effect, if any, of JUMP on pupils' mathematical achievement. The second question was what effects JUMP had on pupils' attitudes to mathematics, while the third was what effect JUMP might have on teacher mathematical knowledge. The fourth question was how well JUMP materials matched with the objectives of the Irish Primary School Mathematics Curriculum (PSMC) and with Irish-developed mathematics materials currently used, and the fifth was how faithfully participating classes implemented the programme.

Has mathematics achievement changed?

In terms of mathematics achievement, pupils who experienced the JUMP programme showed statistically significant gains in mean DPMT scores from the results of the September 2013 to the May 2014 tests. However, so too did pupils in the IMPACT programme. JUMP pupils averaged a gain of nearly seven points (almost half a standard deviation), whereas IMPACT pupils averaged a gain of just over five points (one third of a standard deviation). Both gains would be considered not just statistically significant, but also of pedagogical significance. Although the overall increase in mean scores by pupils in JUMP classes was slightly larger than the increase among IMPACT pupils, it was not statistically significantly larger.

In both programmes, slightly higher gains were made by low- rather than high-achieving pupils (defined as those over one standard deviation below or above the mean when tested in September). Low achievers in both programmes improved by about five score points, while high achievers improved by two points in JUMP and three points in IMPACT. The gains made in each group were broadly similar by gender.

In terms of particular aspects of mathematics (i.e., content and processes), pupils in both groups showed generally similar patterns of performance over the course of the year. Slight differences were apparent in end-of-year performance on Data and on Integrating and Connecting. The slight advantage shown by JUMP pupils on these areas can largely be attributed to three items on which JUMP pupils did *much* better than did IMPACT pupils. The items require pupils to interpret a pictogram in which one symbol represents multiple objects. While the PSMC implies that pupils should develop this skill, the three Irish Third class textbooks reviewed for this evaluation address pictograms only briefly, and limit themselves to examples in which one symbol represents *one* object.[1] However, JUMP dedicates three lesson units to pictograms (called "pictographs" in its materials), and includes examples in which a single symbol represents more than one object. In other words, JUMP satisfies the PSMC aim

[1] As IMPACT materials did not cover the Data strand directly, it seems likely that many IMPACT teachers would have used their usual mathematics textbooks when covering Data topics.

that pupils be able to "read and interpret pictograms" more thoroughly than most Irish textbooks, and this appears to have been borne out in test results.

In sum, the evaluation failed to find a statistically significant effect for JUMP over and above the effect found for IMPACT. Closer analyses of performance did not reveal any particular sub-groups of pupils for whom JUMP was particularly effective, suggesting JUMP is reasonably successful for all, but not particularly successful for any one sub-group of pupils.

Did pupil attitudes change?

Key elements of the JUMP philosophy are that every pupil can be good at mathematics, that teachers should try to decrease pupils' anxiety about mathematics, and that pupils should come to appreciate and be excited by the beauty of mathematics. Therefore, the evaluation examined pupil attitudes to mathematics at the start and end of the year. In most cases, large change in pupil views was not observed. For example, there was no more than a percentage or two difference in those agreeing with these statements: *I wish I didn't have to study maths/I am good at maths/I worry that I won't be able to answer questions in maths class.* There were, however, increases in the percentages of pupils in JUMP classes indicating that they liked maths (6%) and who learned interesting things in maths lessons (8%). The percentage who believed everyone could be good at maths increased by 5% (but increased by 7% among IMPACT pupils).

In JUMP, boys seemed to show a greater increase than girls in positivity towards mathematics, and in mathematical confidence, while the reverse was true of IMPACT. For instance, over the course of the year, there were drops in the percentages of JUMP boys and IMPACT girls who wished they didn't have to study maths. Also, there was an increase of 4% from September 2013 to May 2014 in JUMP boys and IMPACT girls who believed they were good at maths (yet percentages decreased for JUMP girls and IMPACT boys). Although there was little change in the overall percentages of pupils who worried that they would be unable to answer questions in maths class, slightly fewer JUMP boys expressed anxiety about being questioned by the end of the year. Also, the nature of the relationship between questions and mathematics achievement changed slightly: for JUMP, anxiety about being questioned was more weakly correlated with achievement by the end of the year, while for IMPACT, it was much more strongly correlated with achievement than at the outset. In other words, low-achieving IMPACT pupils were more anxious about being questioned in class by May than they had been in September.

As part of the questionnaires, pupils were asked about typical behaviour in mathematics lessons. For the most part, there were few notable changes in responses from September to May or when compared with pupils in IMPACT classes. Regular use of memorisation increased only marginally (by 3%) amongst JUMP pupils during the year, and there was a somewhat surprising larger increase evident amongst IMPACT pupils. However, JUMP pupils were more likely – as would be predicted – to report regularly working on a problem on their own in class.

Among pupils interviewed, the percentage of JUMP pupils who said that mathematics was their favourite subject dropped from 57% (in December 2013) to only 29% (in May 2014). While the drop is quite marked, an almost identical drop was observed amongst IMPACT pupils, suggesting factors external to the programmes may be at play. By the end of the year, pupils in JUMP classes were far more likely to mention (unprompted) that they enjoyed being challenged or stretched in their mathematics lessons. Also, JUMP pupils were overwhelmingly positive about enjoying mathematics and preferring their current mathematics lessons to those in Second class (in contrast to a more ambivalent attitude among IMPACT pupils).

The strength of the correlation between pupils' general positivity to mathematics and mathematics achievement decreased for JUMP pupils over the course of the evaluation, although it increased for IMPACT pupils. In particular, JUMP pupils with above-average achievement scores, and whose scores had improved during the year, were likely to report mildly negative attitudes to mathematics in May. It is possible that the step-by-step, repetitive aspects of the programme may not have appealed to high achievers, or that insufficient use was made of bonus questions for such pupils. However, results should be interpreted with caution due to the small number of pupils involved.

In sum, pupil attitudes and behaviour did change slightly but not as much as might be expected. JUMP did not lead to any notable reduction in pupil anxiety about mathematics, or improved mathematical self-confidence, but mathematics as a school subject does seem to have become slightly more appealing, and, as would be predicted, pupils spent more time engaged in solo work. Differences by gender as well as programme *might* indicate that JUMP promoted positive attitudinal change more effectively for boys than girls, while IMPACT did this more effectively for girls than boys. However, small sample numbers mean that further research would be required to investigate this more fully.

Has teacher knowledge changed?

The evaluation also examined teacher knowledge about teaching mathematics, using a shortened version of the Mathematical Knowledge for Teaching Questionnaire (Delaney et al., 2008; Hill et al., 2004). Baseline MKTQ-S scores collected in September were closely matched for teachers in both programmes, and were slightly above the mean of an Irish norm group of teachers. By the second administration, average number of correct answers increased for teachers in both groups, although very slightly more in JUMP. There was, however, considerable variation with each group in the extent of gains or drops on the MKTQ-S.

Overall, the average improvement among teachers in the JUMP programme was not markedly different from those in the IMPACT programme. Thus, both programmes may have had a positive effect on teacher mathematical knowledge for teaching or the participation in an evaluation of two mathematics programmes may have focused attention on issues related to mathematics teaching (a type of Hawthorne Effect). There may also have been a slight practice effect, although this is unlikely to be marked, given the large amount of time between the two administrations.

In a related vein, an unexpected finding was the noticeable drop in the percentage of JUMP teachers expressing confidence in their ability to adapt their teaching to engage pupil interest. While not directly related to teacher knowledge, knowledgeable teachers tend to be more confident about aspects of their teaching. The reasons for the fall in confidence are unclear. It may be that teachers realised they had heavily relied on JUMP materials and not drawn on their own resources, they may have felt that they did not sufficiently adapt their teaching to pupil need during the year, or there may be other considerations.

How does JUMP align with the PSMC and textbooks?

Most of the teachers who used JUMP felt that it was reasonably well aligned with the Irish PSMC. The most common criticisms related to the use of Canadian expressions and currency, although differences in strand emphases were also raised. The "pitch" of JUMP materials was raised in initial interviews (some teachers thought it too high, others too low) but by only one teacher in the May interviews.

Based on a more formal content analysis of JUMP materials (teacher manuals, pupil materials), it is clear that while some terminology needs adaptation for use in an Irish context, overall JUMP and PSMC content and pitch is quite similar. Generally, the five strands of the JUMP materials correspond closely with those of the Irish PSMC, although some topics appeared under different strands in JUMP. Of the 70 content objectives of the PSMC for Third class, 63 were fully addressed by JUMP materials, six were partly addressed, and only one not addressed at all. The sole objective not addressed related to developing an understanding of the relationship between fractions and division, and is also rarely addressed in Irish-produced materials.

Perhaps surprisingly, the JUMP pupil workbooks were more closely aligned with the PSMC (in terms of relative strand emphasis) than were any of the three Irish textbooks considered. Indeed, the Irish mathematics textbook most commonly used was the one most poorly aligned with the PSMC. Despite this, some teachers felt aspects of JUMP did not align well with the curriculum. For example, there was an inaccurate perception that JUMP paid insufficient attention to Number. This may suggest that some teachers consider their usual mathematics textbooks (some of which over-emphasise Number) as the de facto curriculum.

In sum, despite superficial differences in language and terminology, and some difference in how topics are classified, JUMP represents a reasonably good match with the Irish PSMC.

Was JUMP implemented properly?

JUMP implementation can be examined from the perspective of those who observed lessons, from questionnaire data, and from teachers' own views. Observation data suggest most teachers demonstrated *some* adherence to the relevant lesson plan or programme principles, but few demonstrated very close adherence. Teacher-led instruction and solo work was, as would be expected, more prominent in JUMP than IMPACT, but not markedly prominent vis à vis what might be expected generally. (IMPACT is based on a social constructivist model whereby pupils construct meaning as a group, so teacher-led instruction and solo work should be less common than in mathematics classes *generally*.) The observations suggested that use of some aspects of JUMP (e.g., bonus questions, use of memorisation and repeat procedures, use of workbooks) might have tailed off during the course of the year. Further, the subject matter experts who rated the recorded lessons felt that teachers sometimes missed the spirit of the JUMP programme (e.g., considerable repetition and practice, but no reference to the larger mathematical ideas behind the repeated steps).

From the teacher perspective, most recognised JUMP's positive aspects – the quality and depth of materials, the "ready-made" lesson plans for teachers, the hands-on methods, and the emphasis on building pupil confidence. The main criticism related to the need to adapt the language for an Irish context, the amount of material, and, specific to the present evaluation, there was considerable dissatisfaction about organisational issues.[2] Nearly half of JUMP teachers did not attend the initial CPD day that introduced teachers to the main features of the programme they were to implement during the year. This may partly explain complaints about the amount of material,[3] and why only a minority of teachers used the Confidence Building

[2] As noted in previous chapters, an issue with part of project funding contributed to a delay in project initiation. This in turn caused some organisational problems, most notably, relating to the provision of initial CPD.

[3] Pupils are not expected to complete every part of the JUMP workbook. Some materials are to provide practice opportunities for pupils struggling to grasp a concept, while others are to maintain the engagement of high-achieving pupils. However, this is not immediately obvious from the materials alone.

Unit as intended, while some did not use it at all. Teachers generally expressed dissatisfaction with the amount of CPD provided, the very short notice given for the first two sessions, the webinar format of later CPD, and delays in getting JUMP materials.

CPD is more likely to be found effective by teachers, and to influence their classroom practice, when it is of substantial duration, both in terms of contact hours and overall time span (e.g., Garet, Porter, Desimone, Birman, & Yoon, 2001). Indeed, Guskey and Yoon's (2009) review of the literature found that in order to produce positive effects, CPD initiatives needed to involve 30 or more contact hours, with sustained and structured follow-up. In the present evaluation, the lack of initial training for many teachers, coupled with the paucity of ongoing training (and the technical problems with at least one CPD session) made it harder for teachers to implement JUMP as it should be implemented. Further, anecdotal evidence suggests that JUMP suffered by association with organisational problems involved in its pilot rollout. This may have contributed to reluctance in some parts to fully adopt JUMP methods.

In sum, implementation was not as good as might be hoped, and was poor in some cases. The initial organisational difficulties and poor uptake of CPD could be remedied by a longer run-in time between securing funding and project initiation. However, our understanding is that the relatively limited amount of CPD provided was intentional, based on advice from JUMP staff about what represented adequate support for teachers. Those who provided CPD for IMPACT matched the amount to that provided to JUMP teachers, but noted that it fell well short of what was normally offered. The outcomes suggest that the amount of support provided during the evaluation was insufficient to bring about major change in classroom practice, and that what constitutes adequate levels should be reconsidered.

What contributed to improved achievement?

As noted, the average mathematics achievement scores of pupils in both programmes improved over the course of the evaluation, by a statistically (and a pedagogically) significant amount. However, JUMP gains were not significantly greater than IMPACT gains. As JUMP and IMPACT are based on quite different approaches to teaching mathematics, we must examine what might be associated with improved test performance.

Effects of being evaluated

As has been noted at a number of points, once people know they are being studied, they often modify their behaviour. Behaviour change can be expressed on a continuum from spending a little more time planning mathematics lessons all the way to practising for the test. As evaluators, we made it very clear to all participating teachers that the programme, not the teacher, was the subject of the evaluation. Nonetheless, we are aware from contact with teachers throughout the year that many also felt their teaching was under the microscope.

It is likely that at least some of the teachers involved in each programme put a little more time and thought into lesson plans, or tried a little harder than normal to motivate and engage pupils in mathematics lessons. We also have anecdotal reports from observers that pupils in a few classes may have developed familiarity with the types of questions asked in the DPMT.

In a real-life environment, it can be difficult to separate out the effects of an intervention from the effects of knowing you are part of a study, hence the widespread use of control groups and placebos. IMPACT was not a placebo, but it is essentially an extension of the philosophy of the PSMC. We cannot provide a precise quantum for the Hawthorne effect, but by using a similar approach for both programmes, we can hope that any Hawthorne gain effects are similar and cancel each other out. This means that the *difference* in gain scores between the two programmes is important. An approach that represented a major improvement over that of the PSMC and its underlying philosophy might be expected to show significantly higher achievement gains. Average JUMP gains were slightly, but not significantly, higher than IMPACT gains. The lack of a significant difference may be attributable to any number of reasons, including patchy implementation of JUMP. Nonetheless, JUMP performance has not yet been proven to be significantly better than performance in a control group, indicating that a Hawthorne effect may have contributed to overall improved performance.

Conclusion: Some of the improvement in mathematics achievement is likely to be related to the effects of being evaluated.

Lesson length

The most recent comparable national data for mathematics lessons at primary schools come from TIMSS 2011 (Eivers & Clerkin, 2012; Mullis, Martin, Foy, & Arora, 2012). Conducted at Fourth class, the TIMSS data for Ireland indicate that in 2011 the typical mathematics lesson lasted about 49 minutes. Similarly, the two most recent cycles of the National Assessments of Mathematics found that mathematics lessons averaged 45 minutes in Second class and 52 minutes in Sixth class (2009 Assessment: Eivers et al., 2010), and 43 minutes in Fourth class (2004 Assessment: Shiel, Surgenor, Close & Millar, 2006). All are shorter than the typical lesson duration for JUMP (56 or 57 minutes, depending on time of year) or for IMPACT (55 to 58 minutes). Thus, for example, pupils in this evaluation received about 30 to 40 minutes extra mathematics instruction over the course of an average week than did the Fourth class pupils in TIMSS 2011.

It may be the case that the average amount of time devoted to mathematics instruction has increased since 2011, perhaps due to the launch of the Department of Education and Skills' strategy document *Literacy and numeracy for learning and life* (2011). It may equally be the case that, given their participation in an evaluation of two mathematics programmes, teachers allocated more time than they normally would to mathematics lessons. Whatever the reason for the change, pupils in both programmes are likely to have spent more time in mathematics lessons than did pupils in classes when the DPMT was standardised.

Conclusion: Some of the improvement in mathematics achievement may be related to the extra time allocated to mathematics lessons.

Materials and methods

JUMP materials comprise the Confidence-Building Unit (CBU), the pupil workbooks, and the teacher manuals. Teacher reports show that the CBU was not widely used, or not used appropriately, and its function was not fully understood by most teachers. As such, its efficacy has not really been evaluated here. However, analyses suggest that other JUMP teacher and pupil materials are (despite superficial differences) a reasonably good match for the PSMC. Indeed, in many ways JUMP represents a better PSMC match than the textbooks commonly used in Irish classrooms.

JUMP's simplified design for the pupil Assessment and Practice Workbooks lends itself to a relatively inexpensive workbook designed to be used by pupils, not a glossy book from which problems are copied and completed elsewhere. This was mentioned as a positive by many of the pupils interviewed, who commented on time saved by not having to copy, how copying out was annoying, and how, generally, they liked their workbook much more than the previous year's textbooks.

JUMP materials present "ready-made" lesson plans for teachers. While some might feel that this underplays the importance of the teacher's role in developing and adapting material to meet the needs of their pupils, teachers are able to adapt JUMP material to local need. Indeed, the (Canadian) Pacific Institute for the Mathematical Sciences (2011) has criticised many of the Canadian textbooks used in mathematics lessons as not presenting material in a clear manner, or helping teachers understand the concepts being taught, thereby making it harder for pupils to make connections between content and underlying mathematical concepts. This criticism applies less to JUMP than to some Irish materials, which might benefit from greater clarity in organisation and more guidance for teachers.

Although IMPACT did not provide pupil materials, responses suggest that IMPACT teachers relied less than normal on textbooks, but probably drew on textbooks for the Data strand (for which IMPACT manuals were not available). Data was the only content strand on which there was difference in the overall percent correct scores between the two groups at the end of the year, suggesting that the Data strand might be better covered in JUMP materials than in Irish textbooks. The process area of Integrate and Connect also showed a slight advantage for JUMP pupils, perhaps reflecting the findings in Chapter 3 that JUMP asked pupils to reason, connect and problem-solve, whereas the comparison Irish textbook did not.

Conclusion: Due to limited or incorrect use, it is not possible to judge if the JUMP CBU is effective. However, JUMP pupil and teacher materials appeared to be at least as good as Irish mathematics textbook series currently in use. JUMP's better coverage of aspects of Data, and of Integrating and Connecting, may be related to slightly better performance in these areas.

Future possibilities

To a certain extent, the present evaluation was an incomplete evaluation of the merits of JUMP, and makes it difficult to adjudicate on JUMP's effectiveness. Due to very limited notice, there were poor levels of initial participation in professional development intended to explain JUMP methods and hone skills based on applying JUMP methods. Teacher dissatisfaction with how the initial phases of the evaluation were organised may also have contributed the some cynicism about the *programme's* likely effectiveness.

In addition, the amount of CPD and support provided for the duration of the evaluation was relatively limited. Many teachers complained about feeling isolated and uncomfortable because they spent a few months not only being unsure if they were implementing JUMP correctly, but not knowing how to check or who to ask if they were correct. A consistent research finding is that, almost irrespective of programme efficacy, teachers require ongoing support if it is to be adopted and implemented effectively (e.g., Guskey & Yoon, 2009). This was not the case here, and data from the classroom observations, from teacher self-reports and from pupil descriptions suggest that although there were differences between what typically happened in classrooms in each programme, the differences were not of the size that might be expected.

The fact that JUMP was not fully implemented in many or most classes in this evaluation makes it difficult to establish what effects the programme might have if used as intended. That aside, the programme seems to have merit, it is reasonably popular with teachers and pupils, pupils who took part in JUMP showed achievement gains, but those gains were not significantly greater than gains by pupils in the IMPACT programme. It would be worthwhile to see how JUMP would work if all teachers attended CPD and received more ongoing support. As noted in previous chapters, the short notice given for CPD was related to an unforeseen delay in the release of some project funding, and is highly unlikely to recur. However, the amount and nature of CPD required was agreed with JUMP staff prior to the evaluation, and needs reconsideration, as it appears insufficient to effect behaviour change.

The significant improvements in pupil scores in both JUMP and IMPACT suggest that a further evaluation of both programmes might be merited. If such an evaluation were to be considered, some changes should be made, to the organisation of the evaluation study itself and to JUMP and IMPACT content. The possibility of combining aspects of both programmes and of further analysis of the match between textbooks and the PSMC might also be considered.

Changes to the study

We suggest that any future studies differ from the present one in three main ways. First, and most importantly, teachers should have a much clearer idea from the outset about how JUMP and IMPACT are implemented in practice. Initial CPD should be detailed, precede the adoption of the programme, and be supplemented by more comprehensive training and support during the implementation.

Second, baseline measures of achievement and attitudes should be taken in the spring *before* initial teacher CPD and programme adoption (i.e., in the spring of Second class, assuming that the evaluation focused again on Third class pupils). This would eliminate the possible interference of summer learning loss and slight programme effects in the baseline data, allowing a comparison of spring results with spring results.

Third, a larger sample should be used, ideally split into four groups. The first additional group might combine the most effective aspects of JUMP and IMPACT in a hybrid programme (described in a later section of this chapter). The fourth group would be an "absolute control group" – i.e., a group of pupils simply following the PSMC as normal. We noted earlier that the time spent on mathematics lessons in this evaluation was higher than that found in previous nationally representative samples. This may be due to programme effects or evaluation effects, but it may reflect a broader national trend (attributable to *Literacy and numeracy for learning and life*). If an absolute control group also showed that more time was devoted to mathematics than when the DPMT was standardised, and also showed significantly higher achievement than the test mean, then the improved test scores might simply be attributable to a recent greater focus on mathematics in primary schools.

Adapting JUMP

JUMP CPD and materials both require adaptation in an Irish context, but would benefit from some more general changes too. Regarding local adaptation, some terms and phrases need adaptation in both teacher and pupil materials. This is particularly evident in relation to money, where the Canadian "loonie" and "toonie" either confused or amused, but there are other "Canadianisms" where the language used may interfere with pupil (or indeed teacher) understanding. All sections of JUMP materials should be reviewed and, where necessary, adapted for Irish use prior to any further rollout of the programme (a point also made in the British context by Aduba, 2007). Also, infrequent contact with JUMP experts unfamiliar with the Irish education system, and working in a different time zone, was less than ideal. A better approach would be to supply intensive training in JUMP methodology to a small group of Irish teachers, who could then support the programme locally.

More generally, the initial JUMP CPD day (September 2013, delivered by JUMP founder, Dr John Mighton) was more research-based and theoretical than expected by most of those who attended. While a research base is important, it is also important to discuss practicalities. It is illustrative that most did not understand how to use the JUMP Confidence Building Unit correctly, nor realised that not every child had to answer every question in the pupil materials. Several teachers indicated that watching the sample lesson on perimeter was useful. A set of such videos would provide teachers with a resource on which they could draw throughout the year. Also, while the lesson units gave clear and concise guidelines for teachers, the general introduction and guidelines in the teacher manuals did not. For example, the quite short Introduction section contains a mixture of methods, theory, descriptions of research, quotes from, and information about, Dr Mighton, and some practical advice on implementing the programme. The Introduction should be re-considered and re-designed so it can be referred back to by teachers when they want to check something or refresh their minds about JUMP principles. It is curious, given JUMP's emphasis on scaffolding, repetition and recapping, that none of these feature in the introduction to the teacher materials.

Adapting IMPACT

Teacher feedback indicated that while the IMPACT approach was generally liked by teachers, the materials could be improved, on both a practical and a theoretical level. Dealing first with practicalities, resources in the manual cannot be photocopied for use by pupils, but have to be re-created by teachers. Also, some of the manipulatives recommended by the manuals were difficult to find. Simply adding resources that can be copied or printed and a pack of manipulatives (or recommending only readily accessible material) would improve usability considerably.

The amount of pedagogical theory in the manuals was not popular with many teachers. However, it should be possible to provide a better combination of theory and practical implementation than is found in the current manuals. In particular, the concept of the learning trajectory proved confusing for some. For example, IMPACT identifies two aspects of the learning trajectory – the mode of representation (concrete/pictorial/abstract) and the concept – but does not always show how to combine these to develop and consolidate concepts and processes.

Also, teachers need more guidance on how to integrate some of the (generally very useful) suggested activities into a lesson. In this regard, some exemplar IMPACT lessons would be of use. Finally, the programme should be extended to cover the full PSMC. Although IMPACT *methods* are intended to be transferable across strands, most teachers did not do this, using IMPACT only to teach the strand units covered in the manuals.

Combining aspects of JUMP and IMPACT

Combining JUMP and IMPACT methodology seems counterintuitive, given that JUMP emphasises scaffolding and guided practice, while the social constructivist IMPACT framework emphasises guided discovery. However, the positive features identified in *both* programmes, by teachers, pupils, and mathematics subject matter experts, suggest that an approach that takes the better aspects of each could be worth considering. A full and proper combination of the programmes' best features would involve considerable effort in developing a framework and materials and might run the risk of being too complex and disparate for teachers to implement. It might also, quite understandably, not appeal to those responsible for JUMP or IMPACT. While a worthy idea, we realise it is unlikely to be realised. However, a hybrid model *is* possible, particularly if the proposed changes to each programme outlined in the preceding sections are made.

It is worth looking at the main strengths of each programme. JUMP strengths relate to its comprehensive coverage of curriculum content, good structuring and scaffolding of content, easily implemented lesson plans, and ready-made materials for pupils. In contrast, IMPACT's strengths are its strong focus on group/collaborative activities and fostering of discussion, and a good balance between lower order (e.g., recall concepts or implement procedures) and higher order (reason, communicate, problem-solve) skills. Teachers who become familiar with and understand each approach are likely to employ the different pedagogical approaches in a flexible manner.

JUMP has extensive coverage of the Number and Algebra strand, while IMPACT covers a sizeable number of strand units, but not all. Garforth's (admittedly very small-scale) 2013 study also suggested that JUMP's treatment of Number was more effective than that found in the British Columbia curriculum. Number is the strand typically covered in the first half of the school year, and it lends itself to a very structured approach – which may be more appealing to children in the earlier part of the school year.

JUMP devotes more coverage to Shape and Space than do most Irish textbooks, but level of coverage falls slightly short of that in the PSMC. In contrast, it is extensively covered in IMPACT, and is a strand that lends itself more to discussion, to use of real life materials and manipulatives (i.e., the IMPACT approach). Data is well covered in JUMP (receiving proportionally more coverage than in any of the Irish textbooks reviewed, and slightly more than in the PSMC), but not yet covered in IMPACT. It was also a strand on which JUMP pupils performed well in the present evaluation.

As such, we suggest that any hybrid model might consider using modified JUMP materials and methods to teach the Number and Algebra strand for the first part of the school year. Shape and Space could be taught using IMPACT approaches and (slightly modified) materials, while Data seems suited to a more JUMP-oriented approach. As JUMP also provides reasonable coverage of Measures, a topic not covered in IMPACT materials, it may be more appropriate to use JUMP methods and materials to teach the strand. All this is of course only a rough outline of what form a hybrid model might take, and (as with the programmes themselves) it is highly likely that teachers will modify methods and materials to suit their own situation.

Aligning Irish mathematics textbooks with the PSMC

An unexpected finding from the present study was that JUMP was in many ways a better match to the content and aims of the Irish PSMC than were the more popular Irish mathematics textbooks. Relative to the curriculum, there was a very heavy emphasis in the Irish materials on Number and on isolated computation, with a concomitant limited coverage of Shape and Space. The reason for the mismatch between some textbooks and the curriculum is unclear.

When the PSMC was developed, the National Council for Curriculum and Assessment (NCCA) supplied mathematics textbook specifications to publishers (NCCA, undated). One element of the specifications indicates that Shape and Space should use materials that can be handled and rotated. Perhaps some publishers felt the strand would therefore not need as much coverage as other strands, as teachers would largely deal with it using resources other than textbooks. However, this does not account for the over-focus on Number or isolated computation, especially as the same NCCA document states that "Closely-written pages of ‘sums’ should be avoided.” (p. 6), and that all the strands should receive a balanced level of emphasis and reflect the integrated nature of the curriculum.

Lewis and Archer (2012) noted that 56% of countries participating in TIMSS 2011 had mandated mathematics textbooks at primary school, while 48% had recommended instructional activities. Ireland has neither. We do not propose a single mandated textbook series, and it seems unlikely that the Department of Education and Skills will initiate a system whereby commercial mathematics textbooks must meet minimum standards before they are approved for use in schools. In the absence of greater national oversight, we suggest that teachers, perhaps with advice from the Inspectorate, need to consider how well textbooks match the intent and content of the curriculum. Publishers may tailor materials to marketplace demand, and may originally have been responding to a perceived demand for a strong emphasis on computation. If there is a perceived demand for textbooks and other resources that pay due attention to all strands of the PSMC and its underlying philosophy, it is likely that publishers will respond appropriately.

References

Aduba, N. (2006). *JUMP mathematics: Lambeth Pilot Programme*. Retrieved February 17th, 2014, from http://jumpmath.org/cms/sites/default/files/JUMP%20Lambeth%202006.pdf.

Aduba, N. (2007). *JUMP mathematics in Lambeth: Evaluation and impact on KS2 national tests*. Retrieved February 17th, 2014, from http://jumpmath.org/cms/sites/default/files/JUMP%20Lambeth%202007.pdf

Aduba, N. (2009). *JUMP mathematics in Lambeth: Impact on KS2 national tests 2009.* Retrieved February 17th, 2014, from http://jumpmath.org/cms/sites/default/files/The%20Impact%20of%20JUMP%20Math%20in%20England%20-%20Lambeth%202009.pdf.

Biswas Mellamphy, N. (2004). *JUMP for joy! The impact of JUMP on student math confidence: A report on research findings 2003-2004.* Retrieved August 25th, 2014, from https://jumpmath.org/cms/sites/default/files/JUMP%20for%20Joy!%20Study%202004.pdf.

Bloom, H.S. (1995). 'Minimum detectable effects: A simple way to report the statistical power of experimental designs'. *Evaluation Review*, 19, 5, 547-556.

Charalambous, C., Delaney, S., Hsu, H., & Mesa, V. (2010). A comparative analysis of the addition and subtraction of fractions in textbooks from three countries. *Mathematical Thinking and Learning, 12*, 117-151.

Cooper, H., Nye, B., Charlton, K., Lindsay, J., & Greathouse, S. (1996). The effects of summer vacation on achievement test scores: A narrative and meta-analytic review. *Review of Educational Research, 66 (3)*, 227-268.

Cosgrove, J. (2005). *Issues in the interpretation of PISA in Ireland: A study of the content and design of PISA with comparative analyses of the Junior Certificate Examinations and TIMSS.* PhD diss., National University of Ireland.

Delaney, S. (2012). A validation study of the use of mathematical knowledge for teaching measures in Ireland. *ZDM International Journal on Mathematics Education, 44*, 427-441.

Delaney, S., Ball, D., Hill, H.C., Schilling, S.G., & Zopf, D. (2008). "Mathematical knowledge for teaching": Adapting U.S. measures for use in Ireland. *Journal of Mathematics Teacher Education, 11*, 171-197.

DES (Department of Education and Skills). (2011). *Literacy and numeracy for learning and life: The national strategy to improve literacy and numeracy among children and young people 2011-2020.* Dublin: Author.

Eivers, E. & Clerkin, A. (2012). *PIRLS & TIMSS 2011: Reading, mathematics and science outcomes for Ireland.* Dublin: Educational Research Centre.

Eivers, E., Close, S., Shiel, G., Millar, D., Clerkin, A., Gilleece, L., & Kiniry, J. (2010). *The 2009 National Assessments of Mathematics and English Reading*. Dublin: Department of Education and Skills.

Garet, M., Porter, A., Desimone, L., Birman, B., & Yoon, K. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal, 38 (4)*, 915-945.

Garforth, K.C. (2013). *JUMP Math in a Grade Four classroom.* MA diss, University of British Columbia.

Guskey, T. & Yoon, K. (2009). What works in professional development? *Phi Delta Kappan, 90 (7)*, 495-500.

Hill, H.C., Schilling, S.G, & Ball, D.L. (2004). Developing measures of teachers' mathematics knowledge for teaching. *The Elementary School Journal, 105*, 11-30.

Hughes, K. (2004). *The effects of the JUMP program on elementary students' math confidence.* MA diss, University of Toronto.

Hutchinson, D. and Styles, B. (2010). *A guide to running randomised controlled trials for educational researchers.* Slough: NFER.

Kirk, R.E. (1968). *Experimental design: Procedures for the behavioral sciences.* Belmont, CA: Brooks/Cole.

Learning Mathematics for Teaching Project (2011). Measuring the mathematical quality of instruction. *Journal of Mathematics Teacher Education, 14*, 25-47.

Lewis, M. & Archer, P. (2013). Features of policy and provision. In E. Eivers & A. Clerkin (Eds.). *National Schools, international contexts: Beyond the PIRLS and TIMSS test results* (pp. 13-31). Dublin: Educational Research Centre.

Maciejewski, W. (2012). A college-level foundational mathematics course: evaluation, challenges, and future directions. *Adults Learning Mathematics: An International Journa*l, *7 (1)*, 20-31.

Mesa, V. (2004). Characterizing practices associated with functions in middle school textbooks: An empirical approach. *Educational Studies in Mathematics*, 56, 255-286.

Mighton, J. (2003). *The myth of ability: Nurturing mathematical talent in every child.* New York: Walker & Company.

Mighton, J. (2007). *The end of ignorance: Multiplying our human potential.* Toronto: A.A. Knopf.

Mighton, J. (2013a). *Confidence Building Unit: Fractions Challenge – Level C, D. JUMP Math teacher's manual, second edition.* Toronto: JUMP Math.

Mighton, J. (2013b). *JUMP Math: Fractions Challenge, Level C.* Toronto: JUMP Math.

Mighton, J. (2013c). For the love of math. *Scientific American Mind*, *24 (4)*, 60-67.

Mighton, J., Sabourin, S., & Klebanov, A. (2009). *JUMP Math, book 3 (parts 3.1 and 3.2).* Toronto: JUMP Math.

Mighton, J., Sabourin, S., & Klebanov, A. (2010). *JUMP Math teacher's guide: workbook 3 (parts 3.1 and 3.2).* Toronto: JUMP Math.

Mullis, I.V.S., Martin, M.O., Foy, P., & Arora, A. (2012). *TIMSS 2011 international results in mathematics.* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Murray, B. (2013). *Increased math achievement in Grade 3 and 6 students participating in JUMP Math's 2011-12 National Book Fund Program*. Retrieved August 25th, 2014, from http://www.jumpmath.org/cms/sites/default/files/Student%20Achievement%20From%202011-12%20JUMP%20Math%20Book%20Fund%20(2013).pdf.

National Council for Curriculum and Assessment (undated). *Specifications for publishers: Mathematics*. Dublin: Author.

National Council for Curriculum and Assessment (NCCA) & Department of Education and Skills (1999). *Mathematics curriculum*. Dublin: The Stationery Office.

Pacific Institute for the Mathematical Sciences (2011). PIMS and mathematics education in British Columbia. Vancouver, BC: Dubiel, M.

Remillard, J.T. (2005). Examining key concepts in research on teachers' use of mathematics curricula. *Review of Educational Research, 75*, 211-246.

Shiel, G., Surgenor, P., Close, S., & Millar, D. (2006). *The 2004 National Assessment of Mathematics Achievement.* Dublin: Educational Research Centre.

Stein, M.K., Smith, M.S., Henningsen, M., & Silver, E.A. (2000). *Implementing standards-based mathematics instruction: a casebook for professional development*. New York: Teachers College Press.

Solomon, T., Martinussen, R., Dupuis, A., Gervan, S., Chaban, P., Tannock, R., Ferguson, B. (2011). *Investigation of a Cognitive Science Based Approach to Mathematics Instruction*, peer-reviewed data presented at the Society for Research in Child Development Biennial Meeting, Montreal, March 31 - April 2, 2011.

Tarr, E.J., Reys, R.E., Reys, B.J., Chávez, O., Shih, J. & Osterlind, S.J. (2008). The impact of middle-grades mathematics curricula and the classroom learning environment on student achievement. *Journal for Research in Mathematics Education, 39 (3)*, 247-280.