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.