Hodgen, Jeremy and Coe, Rob and Brown, Margaret and Küchemann, Dieter (2014) Improving students' understanding of algebra and multiplicative reasoning: did the ICCAMS intervention work? In: 8th British Congress of Mathematics Education, 14 - 17 April 2014, University of Nottingham.

Improving students' understanding of algebra and multiplicative reasoning: Did the ICCAMS intervention work?

Jeremy Hodgen, Rob Coe*, Margaret Brown and Dietmar Küchemann
*King's College London, Durham University**

In this paper we report on the intervention phase of an ESRC-funded project, Increasing Competence and Confidence in Algebra and Multiplicative Structures (ICCAMS). The intervention was designed to enable teachers to use formative assessment in mathematics classrooms by evaluating what students already knew, then adapting their teaching to students' learning needs. A key feature was the use of models and representations, such as the Cartesian graph, both to help students better understand mathematical ideas and to help teachers appreciate students' difficulties. Twenty-two teachers and their Year 8 classes from 11 schools took part in the intervention during 2010/11. Pre- and post-tests in algebra, decimals and ratio were administered to the students of these classes, and compared to a control group of students matched from the ICCAMS national longitudinal survey (using propensity score matching). The students in the intervention group made greater progress than the matched control.

Keywords: Algebra, Multiplicative reasoning, Formative assessment

Introduction

Increasing Competence and Confidence in Algebra and Multiplicative Structures (ICCAMS) was a 4½ year project funded by the Economic and Social Research Council in the UK.[1] Phase 1 consisted of a survey of 11-14 years olds' understandings of algebra and multiplicative reasoning, and their attitudes to mathematics (Hodgen et al, 2010). This survey involved both cross-sectional and longitudinal samples. Phase 2 was a collaborative research study with a group of teachers that aimed to improve students' attainment and attitudes in these two areas (Brown, Hodgen and Küchemann, 2012). In this paper, we report on Phase 3 of the study in which the intervention developed in Phase 2 was implemented with a wider group of teachers and students. This paper provides on overview of the intervention and results targeted at the BCME audience. A full consideration of the results will be the subject of a longer paper.

ICCAMS was funded as part of a wider initiative aimed at increasing participation in STEM subjects in the later years of secondary school and university. Our research team at King's College London, having considered the existing research on participation in mathematics (e.g. Matthews and Pepper, 2007; Brown, Brown and Bibby, 2008), felt that the main obstacles to participation lay in negative student attitudes; most students did not want to carry on with their mathematical studies because they believed they were not 'good at mathematics', and 'did not understand it'. They also found it 'boring' and 'unrelated to real life'.

Two mathematical areas which are a key part of the age 11-14 curriculum but which seemed to cause particular problems to students were algebra and multiplicative thinking (ratio and the multiplicative use of rational numbers). Algebra, although not perceived as useful by most students and adults, is particularly important

in relation to further study in mathematics and in subjects that draw heavily on mathematical modelling. Multiplicative thinking is central not only in mathematics but in the application of mathematics in employment and everyday life, especially using percentages and proportions.

Hence, ICCAMS aimed at increasing student participation through improving their understanding of these topics, and, through this, their confidence in their ability to do mathematics. Additionally it also aimed at demonstrating the importance and power of mathematics and its real-life applications.

The ICCAMS approach

Research suggests formative assessment is an effective approach to increasing attainment and engagement (e.g. Black and Wiliam, 1998). Yet, despite widespread take-up of formative assessment nationally and internationally, there is evidence that teachers have considerable difficulties implementing these ideas (e.g. Smith and Gorard, 2005). This may be because formative assessment has been described vaguely and is thus difficult to implement (Bennett, 2011). It may also be because formative assessment has largely been described generically rather than in subject-specific terms (Watson, 2006). Teachers' ability to use formative assessment in mathematics is limited by their knowledge about key ideas, and the likely progression of student learning in them. Thus if teachers focus on teaching mathematical procedures they may find it difficult to see what is causing problems for students in mastering and applying these, and may thus have difficulty responding to the students' difficulties (Hodgen, 2007; Watson, 2006).

In order to address these issues and provide a 'better' and more didactic description of formative assessment, the ICCAMS team drew on the extensive research literature about developing thinking in multiplicative reasoning (e.g. Confrey et al., 2009; Harel and Confrey, 1994) and algebra (e.g. Mason et al., 2005; Watson, 2009). In addition, we developed a set of design principles for which there is research evidence to indicate they are effective in raising attainment (Brown, Hodgen and Küchemann, 2012). These included connectionist teaching (e.g. Askew et al., 1997; Swan, 2006), collaborative work (e.g. Slavin et al., 2009; Hattie, 2009) and the use of multiple representations (e.g. Streefland, 1993; Gravemeijer, 1999; Swan, 2008). In particular, multiple representations, such as the Cartesian graph and the double number line (see, e.g. Küchemann, Hodgen and Brown, 2011), are used both to help students better understand and connect mathematical ideas and to help teachers appreciate students' difficulties.

The ICCAMS teaching materials

The final set of teaching materials consisted of 20 whole class assessment 'starter' activities and 40 lessons.[1] Each assessment starter was designed to inform a pair of linked lessons. For example, the first algebra starter and lesson-pair address the concept of variable and the notion that letters, and expressions involving letters, can represent a range of values simultaneously. The starter, *Which is larger, 3n or n+3?*, is intended to be used some time before the lessons to allow the teacher time to consider the students' approaches prior to teaching. In each of the two lessons, two linear expressions are compared by considering different representations, first, in the context of a boat hire problem, and, second, by returning to the 'pure' context of $3n$ and $n+3$. In both lessons, students are asked to construct Cartesian graphs of the two expressions and then compare these to tabular, word and symbolic representations of

the expressions. Lesson notes provide a description of the lesson together with background materials. The first two pages of the *Boat Hire* lesson are attached as Appendix 1.

Methods

The intervention

The main Phase 3 intervention took place over the academic year 2010/11. Twenty-two Year 8 classes from 11 schools in Hampshire and London took part. Although schools and teachers volunteered to take part in the intervention, the sample of schools included a range of high and low attaining schools. The participating classes were not specially chosen and were the Year 8 classes allocated to the participating teachers. Teachers were asked to use the ICCAMS materials as an alternative to their ordinary teaching of algebra and multiplicative reasoning. Most teachers taught around half of the materials. Teachers also attended six whole day professional development sessions during the academic year led by Hodgen, Küchemann and Brown.

The tests

Tests in algebra, decimals and ratio were administered to the intervention students as a pre-test in October 2010 and again as a post-test in July 2011.[2] The tests were first used in the Concepts in Secondary Mathematics and Science (CSMS) study in the 1970s (Hart et al., 1981). The tests were designed to assess students' conceptual understanding. The algebra test, for example, is designed to assess students understanding of variables.

The intervention focused on the topics more broadly both to ensure the topics covered related to the algebra and multiplicative reasoning topics within the curriculum more generally and to avoid 'teaching to the test'. So, for example, although a key focus within the algebra strand of the intervention was on the use of the Cartesian graph to develop understanding, the Cartesian graph does not feature on the algebra test or either of the other tests. Hence, the tests can be considered to assess the impact of the intervention on students' understanding more broadly.

The sample: Identifying a matched control group

Just over 600 students took part in the Phase 3 intervention. Not all of these students took both the pre- and post-test in all of the tests, although most students took pre- and post-tests in at least one of the tests.[3] The same tests were used in the Phase 1 longitudinal survey. This survey was administered in July 2008, 2009, 2010 and 2011 to a total of 912 students from six schools (including non-intervention classes from the Phase 2 schools). This survey had a dual purpose: it acted as a comparison for the intervention, and was also used to track students' progression across Key Stage 3. As a result, there was no matched control or comparison group in the design. In order to deal with this problem, we used propensity score matching (PSM) to construct matched comparison groups, based on pre-test score and age at pre-test.

PSM is a statistical method first developed by Rosenbaum and Rubin (1983) for estimating causal effects in studies without random allocation. It enables a comparison to be made between treatment and 'control' groups that is based on subsets of both that are well matched on a number of observed characteristics. Under

many conditions PSM can achieve comparison groups that have almost identical distributions on a number of variables simultaneously. This can potentially reduce the problems of trying to interpret differences in outcome measures between two groups that were initially quite different. For this analysis, PSM was set up using logistic regression to predict group membership (intervention/comparison) using pre-test score and age at pre-test as predictor variables. The sample sizes compared in the PSM analysis are shown in Table 1.

Algebra	Number	Ratio
282	292	306

Table 1: Sample sizes for propensity score matching

Results

In this paper we report, first, overall changes in mean score and age, and, second, rates of score gain per year. Other methods, including the use of regression analysis, were also used, but are not reported here. These gave similar results to those reported here, and are discussed more fully in Coe and Hodgen (2012).

Table 2 and Figure 1 show a comparison of the mean pre- and post-tests scores by mean age for both groups (intervention / comparison) in each of the tests (Algebra, Decimals, Ratio)[4].

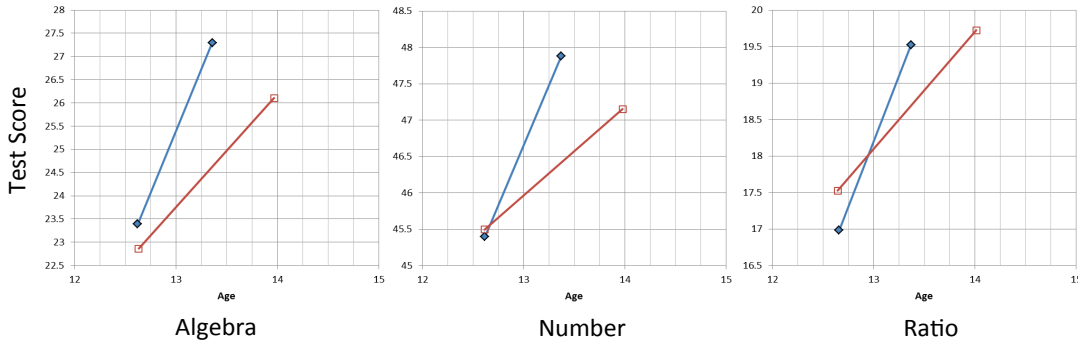


Figure 1: Mean pre- and post scores by mean age for intervention students (Blue) and comparison group (Red) based on propensity score matching for all three tests.

Test	Group	N	Pre-test score	SD	Post-test score	SD	Pre-test Age	SD	Post-test Age	SD
Algebra	Intervention	282	23.39	11.42	27.29	12.74	12.62	0.31	13.96	0.30
	Comparison	282	22.86	12.02	26.10	12.70	12.63	0.37	13.97	0.59
Number	Intervention	292	45.40	14.07	47.88	14.11	12.62	0.30	13.37	0.29
	Comparison	292	45.49	14.47	47.15	15.71	12.61	0.39	13.98	0.58
Ratio	Intervention	311	16.98	8.87	19.52	10.34	12.66	0.35	13.37	0.30
	Comparison	311	17.52	10.07	19.72	10.77	12.64	0.44	14.01	0.62

Table 2: Mean pre- and post- scores and ages for intervention students and comparison group based on propensity score matching for all three tests.[6]

For the comparisons using propensity score matching, the initial pre-test scores of intervention and comparison groups are, not surprisingly, a good deal better

matched. At the time of pre-test, in all but one of the comparisons, the intervention group test mean is actually just below the mean for the comparison group. By the end of the intervention, in all cases the intervention group mean is above that of the comparison group, despite the fact that the average time between tests is over six months shorter for the former.[5]

Mean growth rates, which were obtained by dividing the mean change in test scores by the mean time interval, are shown in Table 3.

	Pooled SD of raw test scores	Rate of growth of mean scores (marks per year)		Standardised rate of growth	
		Intervention	Comparison	Intervention	Comparison
Algebra	12.37	5.29	2.43	0.43	0.20
Number	15.10	3.29	1.22	0.22	0.08
Ratio	10.43	3.57	1.60	0.34	0.15
Weighted mean		***4.03***	***1.74***	***0.33***	***0.14***

Table 3: Growth rates for both interventions and matched comparison groups with pooled (weighted) mean for the main Phase 3 intervention highlighted. The standardised growth rates are calculated by dividing the growth rate (marks per year) by the pooled standard deviation.

It can be seen from these figures that the main Phase 3 intervention groups increased their test scores by an average rate of 4.0 marks per year, compared with a rate of 1.7 marks per year for the comparison groups. These increases correspond to standardised growth rates of 0.33 and 0.14, respectively. Overall, therefore, pupils in the intervention groups have shown roughly twice the rate of increase in test scores of those in the comparison groups. In other words, students in the intervention groups have made the equivalent of about two years' normal progress in one year.

Discussion and implications

These results indicate a substantial effect for the ICCAMS intervention equivalent to a doubling of the annual rate of learning. This effect is of a similar order to that found for formative assessment (Wiliam et al., 2004). However, a major criticism of that original study was that formative assessment was described largely generically and, thus, is difficult for teachers to implement in mathematics education. The ICCAMS intervention fills this gap by providing support and guidance to enable such implementation. Hence, this study provides further weight to the evidence on formative assessment, although we note that the ICCAMS intervention drew on the mathematics education literature more broadly in its implementation of formative assessment.

It is important to express a number of caveats to these findings. First, the design did not include a matched control or comparison group and, hence, it is likely that there may be important unobserved differences between students in the intervention and the matched groups. In addition, tests for the intervention classes were administered twice in the same year (at the start and end of the year) whereas for most of the comparison group the tests were all taken at the end (June/July) of the academic year. Hence, the comparison group may suffer a disadvantage due to the summer break. We note, however, that, in a previous study of progression in primary school, the rate of learning during the summer appeared to be on a par with that for the rest of the year for Key Stage 2 students (Brown et al., 2008). Second, the teachers involved in the interventions were self-selected volunteers. Third, the intervention,

and in particular the professional development for teachers, was undertaken by the design team.

Hence, whilst these results provide sufficiently strong evidence to justify further evaluation of the ICCAMS intervention, any interpretations of these differences as causal effects of the interventions must be cautious. The results are best interpreted as indicating the need for a further evaluation involving randomised allocation of students and teachers to intervention and control groups.

Endnotes

1. We are grateful to the ESRC for funding this study (Ref: RES-179-34-0001).
2. The full title of the Decimals test is Number 2 (Decimals and Place Value). The Ratio test is titled Test R to avoid indicating the items involve ratio.
3. The numbers of Phase 3 intervention students with both pre- and post-tests are 363 (Algebra), 401 (Decimals) and 399 (Ratio). Pre-tests for one Phase 3 school, involving two classes and approximately 60 students, were not carried out due to time constraints within the school.
4. The maximum scores on the tests are 51 (Algebra), 73 (Decimals) and 24 (Ratio). See Hart et al. (1981) for further information. Note that the year on year gain is relatively small (Hodgen et al., 2010).
5. In all cases, the difference in the growth rates of the two groups is statistically significant (Coe and Hodgen, 2012).
6. The comparison group students tend to be older as a result of the test administration. This is a discussed as a potential limitation later in the paper.

References

Askew, M., Rhodes, V., Brown, M., Wiliam, D & Johnson, D. (1997) *Effective Teachers of Numeracy*. London: King's College London.

Bennett, R.E. (2011) Formative assessment: a critical review. *Assessment in Education: Principles, Policy & Practice,* 18(1), 5-25.

Black, P.J. & Wiliam, D. (1998) Assessment and classroom learning. *Assessment in Education,* 5(1), 7–73.

Brown, M., Askew, M., Hodgen, J., Rhodes, V., Millett, A., Denvir, H. & Wiliam, D. (2008) Individual and cohort progression in learning numeracy ages 5-11: Results from the Leverhulme 5-year longitudinal study. In Dowker, A. (Ed.) *Children's Mathematical Difficulties: Psychology, Neuroscience and Education* (pp. 85-108). Oxford: Elsevier.

Brown, M., Brown, P. & Bibby, T. (2008) "I would rather die": Reasons given by 16 year-olds for not continuing their study of mathematics. *Research in Mathematics Education*, 10(1), 3-18.

Brown, M., Hodgen, J. & Küchemann, D. (2012) Changing the Grade 7 curriculum in algebra and multiplicative thinking at classroom level in response to assessment data. In Sung, J.C. (Ed.), *Proceedings of the 12th International Congress on Mathematical Education (ICME-12)* (pp. 6386-6395). Seoul, Korea: International Mathematics Union.

Coe, R. & Hodgen, J. (2012) *Analysis of differences in test score gains between pupils in intervention and comparator classes. ICCAMS Technical Report 06*. Durham / London: Centre for Evaluation & Monitoring, Durham University / King's College London.

Confrey, J., Maloney, A.P., Nguyen, K.H., Mojica, G. & Myers, M. (2009) Equipartitioning / splitting as a foundation of rational number reasoning using learning trajectories. In Tzekaki, M., Kaldrimidou, M. & Sakonidis, H. (Eds.), *Proceedings of the 33rd Conference of the International Group for the*

Psychology of Mathematics Education (Vol. 1, pp. 345-352). Thessaloniki, Greece: PME.
Gravemeijer, K. (1999) How Emergent Models May Foster the Constitution of Formal Mathematics. *Mathematical Thinking and Learning*, 1(2), 155-177.
Harel, G. & Confrey, J. (1994) *The Development of Multiplicative Reasoning in the Learning of Mathematics*. Albany, NY: State University of New York Press.
Hart, K., Brown, M.L., Küchemann, D.E., Kerslake, D., Ruddock, G. & McCartney, M. (1981) *Children's understanding of mathematics: 11-16*. London: John Murray.
Hattie, J. (2009) *Visible Learning: A synthesis of over 800 meta-analyses relating to achievement*. Abingdon: Routledge.
Hodgen, J. (2007) Formative assessment: tools for transforming school mathematics towards dialogic practice? In Pitta-Pantazi, D. & Philippou, G. (Eds.) *Proceedings of CERME 5: Fifth Congress of the European Society for Research in Mathematics Education, 22 – 26 February* (pp. 1886-1895). Larnaca, Cyprus: European Society for Research in Mathematics Education / Department of Education, University of Cyprus.
Hodgen, J., Brown, M., Küchemann, D. & Coe, R. (2010) Mathematical attainment of English secondary school students: a 30-year comparison. Paper presented at the *British Educational Research Association (BERA) Annual Conference*, University of Warwick.
Küchemann, D.E., Hodgen, J. & Brown, M. (2011) Using the double number line to model multiplication. In Pytlak, M., Rowland, T. & Swoboda T. (Eds.) *Proceedings of the Seventh Congress of the European Society for Research in Mathematics Education (CERME7)* (pp. 326-335). Poland: University of Rzesów.
Mason, J., Graham, A. & Johnston-Wilder, S. (2005) *Developing Thinking in Algebra*. Maidenhead: Open Univ. Press.
Matthews, A. & Pepper, D. (2007) *Evaluation of participation in A level mathematics: Final report*. London: QCA.
Rosenbaum, P.R. & Rubin, D.B. (1983) The central role of the propensity score in observational studies for causal effects. *Biometrika,* 70(1), 41-55.
Slavin, R.E., Lake, C. & Groff, C. (2009) Effective programs in middle and high school mathematics: A best- evidence synthesis. *Review of Educational Research*, 79(2), 839-911.
Smith, E. & Gorard, S. (2005) "They don't give us our marks": the role of formative feedback in student progress. *Assessment in Education: Principles, Policy and Practice, 12*(1), 21-38.
Streefland, L. (1993) Fractions: A Realistic Approach. In Carpenter, T.P., Fennema, E. & Romberg, T.A. (Eds.), *Rational Numbers: An Integration of Research*. Mahwah, NJ: Lawrence Erlbaum.
Sutherland, R., Rojano, T., Bell, A. & Lins, R. (Eds.) (2000) *Perspectives on School Algebra*. Dordrecht: Kluwer.
Swan, M. (2006) *Collaborative Learning in Mathematics: A challenge to our beliefs and practices*. London: NIACE.
Swan, M. (2008) A designer speaks: designing a multiple representation learning experience in secondary algebra. *Educational Designer*, 1(1).
Watson, A. (2006) Some difficulties in informal assessment in mathematics. *Assessment in Education,* 13(3), 289-303.
Watson, A. (2009) Paper 6: Algebraic reasoning. In Nunes, T., Bryant, P. & Watson, A. (Eds.) *Key understandings in mathematics learning*. London: Nuffield Foundation.
Wiliam, D., Lee, C., Harrison, C. & Black, P.J. (2004) Teachers developing assessment for learning: impact on student achievement. *Assessment in Education: Principles, Policy and Practice,* 11(1), 49-65.

Appendix 1: The first two pages of the teaching notes for Lesson 1A, The Boat Hire Problem.

Boat Hire

Olaf is spending the day at a lake.
He wants to hire a rowing boat for some of the time.

Freya's Boat Hire charges £5 per hour.
Polly's Boat Hire charges £10 plus £1 per hour.

Whose boat should Olaf choose?

Summary

In this lesson, the boat hire problem is used to explore the two algebraic relationships underlying Freya's and Polly's different hire charges.

A variety of representations are used to express the relationships:
everyday language,
algebraic expressions,
tables of values,
points on a Cartesian graph.

Outline of the lesson

1. Display the Boat Hire problem and ask students for their immediate responses.
   - Ask students to consider the problem further in small groups.
   - Collect numerical data on the total cost for various numbers of hours (and listen to students' arguments and conclusions - but don't pursue these at this stage).
   - Represent the data
     - 'randomly' on the board
     - in a (randomly ordered) table.
   - Represent the data in an ordered table
     - try to prompt the need for this, rather than simply produce such a table.

2. Ask students to represent the hire-rules as algebraic expressions (eg $5a$ and $10+a$), or algebraic relations (eg $b = 5a$, $b = 10+a$).

3. Ask students to represent the data as points on a standard (Cartesian) graph.
   - Are Freya's and Polly's charges ever equal?

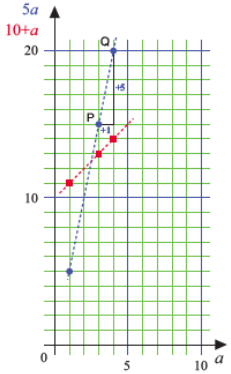
4. Discuss, use, make links between the various representations and between them and the story.

not ordered

a	$5a$
1	5
4	20
3	15

ordered

a	$5a$
1	5
3	15
4	20



Overview

Mathematical ideas

When we work with equations, we often think of a letter as representing a single number as yet unknown. Here, we are working with relations between two variables (the number of hours and the charge) and we think of the letters as representing a set of numbers.

A Cartesian graph is a particularly powerful way of representing variables, since it allows us to represent a range of values simultaneously.

Students' mathematical experiences

Students might discover some of the following

- for some values of a, Freya's hire charge ($5a$) is larger than Polly's ($10+a$), but for others it is smaller
- when $a = 2.5$ the expressions are equal
- if a increases by 1, then $5a$ increases by 5, but $10+a$ only increases by 1

  each set of points on the graph forms a pattern: each lies on a straight line.

Students might discuss

- different slopes and how these relate to the hourly charges
- continuity, ie whether some or all points on the line fit the relationship.

Some students may want to change the scales of the axes. Discourage them from doing so. The expressions have been deliberately chosen to be represented on an equally-scaled graph.

Key questions

When is Freya (or Polly) cheaper?

How could we record this more systematically?

What happens to the cost as the number of hours increases?

What if Olaf hired a boat for 1.5 hours?

Assessment and feedback

Be flexible over the organisation and timing of the lesson. Some teachers have taught this lesson over two periods.

Choose some students to contribute to a subsequent discussion. Some (less confident?) students' contributions may be more coherent if you "rehearse" with them beforehand: "That's a great idea. I'm going to ask you to explain that to the class. Let's have a go at preparing what you'll say."

Allow students time to generate algebraic expressions, but if they really struggle you may want to provide the expressions for them.

Some students may have difficulties constructing a Cartesian graph. Observe the students and decide whether you need to spend some time with either a group or the whole class teaching these skills.

You may need to prompt or challenge some students to consider whether the charges can ever be equal.

Adapting the lesson

You might want to adopt a different context for a subsequent lesson. For example, the price of an ice cream cone for different numbers of scoops, or the yearly cost of belonging to a swimming club (based on a membership fee and a cost for each visit). Choose the numbers carefully - for example, keep the multipliers small if you want to graph the relationships, and think about where the values coincide - do you want this to occur for a simple whole number (4 ice cream scoops, say), or something more obscure (3.4 scoops, say)?