The impact of using Numeracy Ninjas on the numeracy fluency of Key Stage 3 students in a large comprehensive secondary school

William Emeny
Curriculum Leader for Mathematics and Leading Practitioner
Wyvern College, Fair Oak, Hampshire, UK

Significant contributions by
Catherine Bell, Clair Hosie, Charlotte Kay, Helen Lo, Lauren Ross and Frances Walker
Wyvern College, Fair Oak, Hampshire, UK

10 June 2016

Executive Summary

After a three-year professional learning journey we created Numeracy Ninjas, a KS3-focussed numeracy activity to support students becoming highly-fluent in elementary numeracy skills and skills that are the most important foundations for studying GCSE Mathematics.

After launching in summer 2015, Numeracy Ninjas is now estimated to be in use in over 1800 schools worldwide. In order to help schools considering using Numeracy Ninjas answer the question “what impact will it have on my students’ outcomes?”, we conducted a year-long study to evaluate its impact on the numeracy fluency of our Year 7 and 8 students at Wyvern College. The study involved over 470 students at Wyvern College using Numeracy Ninjas.

We found:
1. Numeracy Ninjas increased the average student’s gain in fluency with mental calculation strategies by at least an additional 70% when compared to the gains made by a control cohort\(^1\).

\(^1\) Measured using an in-house developed diagnostic numeracy assessment
2. Numeracy Ninjas increased the average student’s gain in fluency with essential skills required to study GCSE Mathematics\(^2\) by at least an additional 70% when compared to the gains made by a control cohort\(^1\).

3. Numeracy Ninjas had a statistically significant positive impact on the attainment outcomes of our KS2 level 4 students in terms of their fluency with mental calculation strategies, times tables and the essential skills required to study GCSE Mathematics\(^1\).

4. The impact on KS2 level 2 and 3 students was likely positive, but we could not prove statistical significance at \(\alpha = 0.05\) due to relatively small sample sizes.

5. The impact on KS2 level 5 and 6 student’s numeracy fluency and their fluency with the essential skills required to study GCSE Mathematics were inconclusive. We believe there may still be some benefit in using Numeracy Ninjas with these students, but done so more diagnostically and on a lower frequency basis than with KS2 level 4 (and below) students.

6. By sampling students’ performance on Numeracy Ninjas on a weekly basis we believe that we can use Numeracy Ninjas even more effectively in the future. All KS2 level 4 (and below) students will use Numeracy Ninjas every lesson until they regularly achieve Ninja Scores of 30. After approximately week 10 teachers will be encouraged to use Numeracy Ninjas more diagnostically, using Assessment for Learning strategies, to identify which topics students still cannot answer and explicitly teach them how to answer the questions correctly. The diagnostic approach will be used from the start with KS2 level 5 (and above) students running Numeracy Ninjas on a weekly or twice weekly basis until they regularly achieve Ninja Scores of 30.

\(^2\) See ‘Prior Learning Dependencies Network’ section of this report for details
Numeracy Ninjas- three years in the making

Numeracy Ninjas was created from the accumulation of three years of professional learning focusing on how cognitive science research findings, particularly those relating to memory, could be used to improve students' fluency with mathematical concepts that underpin successful GCSE Mathematics study.

The professional learning journey started by researching and identifying the specific mental numeracy strategies and elementary mathematical concepts that are most beneficial to have stored in long-term memory. Through this essential prior knowledge being accessible automatically, and without significant demand on working memory, these skills, facts and concepts then do not become cognitive-resource-consuming barriers to learning the higher-level mathematical concepts which rely upon them.

Once we had identified the skills in which we sought our students to have a high degree of fluency, we then proceeded to develop a diagnostic numeracy assessment to understand our students’ numeracy fluency in year 7 and 8 at Wyvern College. From the resulting analysis, we observed what we considered to be a significant number of issues relating to a lack of fluency with elementary numeracy skills, even with our KS2 highest attaining students.

Subsequently, we studied the literature relating to memory, specifically retention and transfer of learning, to understand how research findings could inform us in our development of a high-impact Key Stage 3 numeracy intervention resource. The aim was for this to result in significant and sustained fluency gains (within the context of the numeracy gaps identified in the diagnostic testing) and ultimately to support enhanced GCSE Mathematics outcomes for the participating students.

Prior learning dependencies network

In 2013, we visited the Mathematics Department at King Solomon Academy, London led by Bruno Reddy. The department had (and still continues to have) a strong reputation for its development and delivery of a ‘depth-before-breadth’, mastery-based mathematics curriculum and the use of cognitive research findings to inform the pedagogical practices of the teachers at the school. During my visit, we met Kris Boulton, who introduced us to the notions of working memory, retrieval and storage strength, the spacing effect and also the benefits of interleaving practice.

In subsequent reading around the concept of working memory and about Cognitive Load Theory (Sweller, 1988), we realised the importance of this idea within the domain of
Specifically, if students' working memories are overloaded by a lack of fluency with the prior-learning skills underpinning the current topic of study, the cognitive resources they have left to devote to processing the learning of this topic are reduced and the learning is sub-optimal. Anecdotally, if you are trying to teach a lesson on factorising algebraic expressions and students need to devote considerable working memory resources to identifying the highest-common-factor of the term coefficients, the resources they have left to think about the concept of algebraic factorisation are reduced. Daniel Willingham (2009) talks about memory being ‘the residue of thought’. If students spend your lessons thinking about the prior-learning skills that are the foundations of your lesson objective, they will not be learning the intended objective.

Following on from this reading, it occurred to us that mathematics teachers may need to consider some learning objectives as more important than others in terms of their students’ fluency. We asked ourselves the question, ‘what are the Key Stage 3 concepts that form the essential prior-learning skills for the most number of topics in GCSE Mathematics?’ Or put another way, ‘which topics are the biggest barriers-to-learning in GCSE Mathematics if students lack fluency in them?’

In 2014 we published an analysis of the GCSE Mathematics curriculum looking at the relationships between the syllabus learning objectives in terms of the prior learning dependencies (Emeny, 2014). We organised the syllabus into 164 different mathematical skills, representing each one as a node. Where you needed to be able to perform one of the skills in order to do another, we joined them with a link. For example, the ‘finding factors’ node was linked to the ‘simplifying fractions’ node because you need to be able to do the former in order to do the latter. We identified 935 prior learning links and used the online freeware, Gephi, to visualise the resulting curriculum map. We scaled the node size by the number of out-going links; i.e. the largest nodes represent the concepts that were the required prior learning for the most number of topics.
We hypothesised that if we could support our Year 7 and 8 students in developing high levels of fluency with the largest nodal topics in the curriculum map, that in subsequent study on the GCSE Mathematics course in Years 9, 10 and 11, these topics would not form working-memory-consuming barriers-to-learning the higher level concepts resulting in better learning of those concepts. Just as a student who had to phonetically sound out words would unlikely be able to appreciate the linguistic flair of Shakespeare, we considered that students who have inefficient or no strategies for elementary mental calculation may have limits imposed on their appreciation and learning of GCSE Mathematics concepts.
Inspired by this analysis, and seeing the consistency of priorities identified by it and those in Bruno Reddy’s mastery curriculum, we redesigned our Key Stage 3 curriculum to ensure we taught the ‘large node’ topics in increased depth.

Diagnostic numeracy testing

To understand extent of the need for this curriculum approach amongst our students at Wyvern College and to be able to assess the impact of the changes we had made, we developed our own diagnostic numeracy assessment.

The assessment has been sat by all Year 7 and 8 students at Wyvern College for the last two years and measures students’ fluency with 25 mental numeracy strategies, a sample of 30 of the times tables up to $10 \times 10$ and 29 of the largest nodal topics (which we called ‘key skills’). The mental numeracy strategies were identified from the Department for Education publication, *Teaching Children to Calculate Mentally* (Department for Education, 2010) and included elementary numeracy skills such as: number bonds to 10, adding near doubles with compensating and understanding when it is beneficial to count from the smallest to the largest number in a subtraction.

The test is administered over a single lesson, taking approximately 50 minutes to complete. Students watch a series of questions on a Powerpoint presentation and have a time limit to choose from 5 multiple-choice answers. Each answer has an associated letter A-E and the students colour in the corresponding circle on a Quick Key optical reader answer sheet. Whilst the time limits vary depending upon the question, they are relatively short to ensure that only students who can identify the correct answers using an efficient strategy will be able to record their answer in time. The intention of the assessment is not simply to find out if students can answer questions on the different strategies, but to find out if they can do so efficiently without placing significant sustained load on their working memories in order to identify the correct solutions.

---

3 Quick Key is a web application that enables teachers to perform multiple-choice questioning efficiently. Students shade in their answers on the optical reader sheet and the teacher then scans the sheet via a smartphone app that marks the assessment.
Figure 2- Example of a Quick Key answer sheet used by students when sitting the diagnostic numeracy assessment

<table>
<thead>
<tr>
<th></th>
<th>KB2 3.0</th>
<th>KB2 3.5</th>
<th>KB2 3.7</th>
<th>KB2 4.0</th>
<th>KB2 4.3</th>
<th>KB2 4.7</th>
<th>KB2 5.0</th>
<th>KB2 5.5</th>
<th>KB2 5.7</th>
<th>KB2 6.0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number bonds to 6 (Q1)</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
</tr>
<tr>
<td>Number bonds to 5 (Q2)</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
</tr>
<tr>
<td>Number bonds to 10 (Q3)</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
</tr>
<tr>
<td>Number bonds to 10 (Q4)</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
</tr>
<tr>
<td>Number bonds to 20 (Q5)</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
</tr>
<tr>
<td>Number bonds to 30 (Q6)</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
</tr>
<tr>
<td>Number bonds to 100 (Q7)</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
<td>10%</td>
</tr>
</tbody>
</table>

Figure 3- Mental numeracy strategies results for the Year 7 cohort at Wyvern College in 2014-2015. The percentages represent the proportion of each Key Stage 2 prior attainment group that got the corresponding numeracy skill correct on the assessment.
The results from the diagnostic testing showed us the prevailing assessment culture at the time, to summarise students’ proficiency by a single national curriculum level, was misleading (Emeny, 2015). Many level 5 and 6 students had level 2 numeracy gaps. In students being awarded a level 5 it could not be assumed that they had fluency will all the below-level-5 topics. Whilst their level 2 numeracy gaps had not limited them in reaching level 5 up until that point, would they in the future prevent or slow the progress those students could make?

Our students’ lack of fluency with some of the most elementary numeracy skills surprised us. For example, when given the question, 2 + 359, only 57% of our KS2 level 4.0 students answered the question correctly having been given 10 seconds to do so. Reversing an addition to make it easier, such a basic strategy, was not a fluent skill for a significant number of our students for whom their KS2 prior attainment would not be considered ‘low’.

We inferred that the diagnostic numeracy assessment results showed that all students, not just the Key Stage 2 low-attaining, needed a regular activity in their Key Stage 3 mathematics lessons that would identify and rectify any lack of fluency with elementary numeracy skills.

Desirable difficulties

We conducted a literature review to understand the concept of fluency better. If we were to create a fluency-building intervention, we needed to understand the definition of fluency within the context of mathematics education and then strategies that impact on it in depth.

However, before we read around the notion of fluency, we learned of the important, recurring theme that appears in educational research on learning; the much discussed distinction between learning and performance. Whilst the jargon associated with the concepts has changed over the past century, the important idea has not. Learning must have elements of retention over time and contextual transferability, whereas performance is just how a learner can perform at an instantaneous point in time. Performance can be measured, learning must be inferred.

Performance is often a poor proxy for learning. Things such as recency of retrieval, mimicry and priming can enhance performance, but we should not necessarily infer from that performance that the knowledge will be sustained over time and/or that the student will be able to apply it to other contexts. We cannot be sure that learning took place.

The New Theory of Disuse (Bjork & Bjork, 1992) gave us a conceptual framework to understand how to design learning activities (and sequences of activities) that are focussed on building retention and transferability of learning over time rather than short-term performance. We learned that the notions of performance and learning could be mapped
onto corresponding notions of retrieval and storage strength related to memory. Rather than seeing forgetting as detrimental, we learned about its importance in enhancing learning. By limiting gains in retrieval strength by using desirable difficulties such as spacing and interleaving (Rohrer & Taylor, The shuffling of mathematics practice problems improves learning, 2007), (Rohrer & Pashler, Increasing retention without increasing study time, 2007), (Rohrer, The effects of spacing and mixing practice problems, 2009) it has been shown that you can enhance gains in storage strength (learning).

We were particularly grateful for, and inspired by, an outstanding contribution (Cepeda, Vul, Rohrer, Wixted, & Pashler, 2008) that suggested the optimum spacing interval is approximately 3 weeks (if the test is a relatively long time away).

We also learned about the benefits of the testing effect (J, Rawson, Marsh, Nathan, & Willingham, 2013) and how regular retrieval practice could enhance long-term retention of learning.

Subsequently, we felt in a more informed position to define fluency within the context with which we were working as learning that is stored in long-term memory with both high retrieval and storage strength. In this case, being fluent with the numeracy and prior-learning skills required to study GCSE Mathematics meant learners needed to be able to retrieve and process the required necessary information with low cognitive load and to be able to apply it to a broad range of different contexts. This kind of learning can only be achieved over time and we realised a structured, regular, planned approach was needed to gain the benefits of the desirable difficulties.

We summarised our learning (Emeny, Forgetting is necessary for learning, desirable difficulties and the need to dissociate learning and performance, 2015) and then began to think about how we could incorporate these ideas into a learning activity that could use them to tackle the numeracy needs our students had that were identified from the diagnostic testing.

Numeracy Ninjas was born.
What is the Numeracy Ninjas initiative?

Numeracy Ninjas is run by teachers, typically in Year 7 mathematics lessons. Each student has their own Ninja Skill Book which features quiz activities containing 30 questions. There are 10 questions on mental strategies, 10 on the timetables and 10 on the key skills.

![Ninjas Skill Book Example](image)

Students have 5 minutes to work on the questions. At the end of the session the teacher shows a Powerpoint presentation containing the answers and the students mark their work. They record their Ninja Score (out of 30). This corresponds to a Ninja Belt colour. Students aspire to achieve a Ninja Score of 30 and acquire the coveted Black Belt.
After the students have marked their work it is suggested that the teacher asks the class if they would like to discuss any of the question they got wrong or couldn’t do. This is meant to be a ‘little and often’ approach- getting better question-by-question, day-by-day.

A full Ninjas session is meant to take no more than 10 minutes at the start of a lesson. This is possible if efficient routines are established.

The topics of the 10 mental strategies and 10 key skills questions rotate on a weekly basis. There are 30 of each respectively which repeat every 3 weeks to support the building the of retention of our students’ learning. Thus, we incorporated into the design of Numeracy Ninjas our learning about the benefits of spacing and interleaving.

The Ninja Skill Books, an answers Powerpoint presentation and other promotional materials were published on the Numeracy Ninjas website, www.numeracyninjas.org, in August 2015, free for any school to download and use for non-profit-making educational purposes.

Additional promotional material such as pencils, badges, stickers and rulers are also available from an online shop on the website.
Impact study design

Cohorts

Our current Year 8 cohort completed the diagnostic testing when they were in Year 7 and so we used this cohort as a control group as Numeracy Ninjas wasn’t created until they the summer before they started Year 8. Whilst in Year 7, 13 of the lowest attaining students took part in an intervention called Numeracy Ambassadors. This was a tutor time activity they participated in 3 times a week where they completed numeracy-focussed activities supported by a Year 10 student (Numeracy Ambassador). The whole cohort is referred to as C1 NA for the remainder of this report. The subsequent analysis is based upon assessment data of 233 students in this cohort.

When the C1 NA cohort reached Year 8 they all participated in Numeracy Ninjas. Every lesson started with the 5-minute Numeracy Ninjas session in their Skill Books. By approximately Christmas time (week 15) classes whose majority of students were regularly attaining a Ninja Score of 30 (full marks) dropped to using Numeracy Ninjas twice per week for the remainder of the year. Numeracy Ambassadors did not run for this cohort when they were in Year 8. This cohort (when they were in Year 8) is referred to as C1 NN for the remainder of this report. The subsequent analysis is based upon assessment data of 248 students in this cohort.

The Year 7 cohort that started at Wyvern College after Numeracy Ninjas was created used it initially in all lessons. As per the C1 NN cohort, high performing classes dropped to using it twice per week from approximately week 15 onwards. This cohort also participated in the Numeracy Ambassadors initiative. This cohort is referred to as C2 NANN for the remainder of this report. The subsequent analysis is based upon assessment data of 241 students in this cohort.

Data collection

All cohorts took the in-house diagnostic numeracy assessment at the start and towards the end of each academic year. The months in which these were taken are stated on subsequent figures in this report.

The Ninja Scores (out of 30) of all students in all classes was recorded on a weekly basis on a shared Google Spreadsheet. The intention of this was to supply us with a continuous proxy for the impact Numeracy Ninjas was having on students’ numeracy fluency as the year progressed and in between the points at which the diagnostic numeracy assessment was taken. It also allowed us to identify our Grand Masters each term, the highest-performing and
most-improved student. The Grand Masters received certificates and prizes including a Numeracy Ninjas pencil, blazer badge and gold Grand Master sticker for their exercise book.
Results and discussion

Diagnostic numeracy assessment results

Figure 6 shows the mean scores for the C1 NA, C1 NN and C2 NANN cohorts on the diagnostic numeracy assessment as a whole and then broken down into the mental strategies, timetables and key skills sections.

Figure 6 - Mean cohort scores of diagnostic numeracy assessment. Error bars show the 95% confidence intervals.

The cohorts who used Numeracy Ninjas, C1 NN and C2 NANN, outperformed the control cohort, C1 NA, who only used Numeracy Ambassadors. The Numeracy Ninjas-led gains in fluency were greater in all three areas of the diagnostic test: mental strategies, timetables and key skills.

The fluency gains were greatest for the C2 NANN cohort. This is understandable, given that the C1 NN cohort started from a position of having made some progress during Year 7 the year before.

Significance testing was carried out using two-tailed student t tests with $\alpha = 0.05$. Hypotheses were tested based on comparing the mean scores of the C2 NANN cohort versus the C1 NA cohort. The following hypotheses were found to be statistically significant:

1. Numeracy Ninjas increased the average student’s result in the mental strategies section of the diagnostic numeracy assessment by at least an additional 70% when compared to the gains made by the C1 NA control cohort.
2. Numeracy Ninjas increased the average student’s result in the key skills section of the diagnostic numeracy assessment by at least an additional 70% when compared to the gains made by the C1 NA control cohort.
The gains the C1 NN cohort made with their timestables were not found to be statistically significant. We believe this was due to a large proportion of the KS2 level 5 and 6 students ‘topping out’ on the timestables section after already in a strong position in June 2015. There was only a small amount of progress to be made before gaining full marks on the timestables section.

The diagnostic numeracy assessment results were then viewed in terms of differing KS2 prior attainment level groups to understand the relative gains each group made using Numeracy Ninjas.

Figure 7 shows the fluency gains made by the KS2 level 2 students in the C2 NANN cohort. Statistical significance cannot be tested as there were no level 2 students in the C1 NA cohort. Whilst conclusions should be approached cautiously with such a small sample size and spread in the results, it was disappointing that the apparent fluency gains in mental strategies and timestables were not matched by similar gains in the key skills. We cautiously pose the following questions for future discussion and research:

1. Do these KS2 level 2 students have other learning difficulties in acquiring fluency with the key skills that are unaffected by improvements in fluency with the mental strategies and timestables?
2. To what degree are the improvements these KS2 level 2 students have made based on understanding rather than fact memorisation?
3. Do they not reach the key skills questions during a 5-minute Ninja session because they are still working on their fluency with the other two areas. I.e. are they running out of time before reaching the key skills section?

Figure 8 shows the fluency gains made by the KS2 level 3 students in the C1 NA and C2 NANN cohorts.

![KS2 level 3 students' scores on the diagnostic numeracy assessment](image)

**Figure 8**- Key Stage 2 level 3 students’ mean scores on the diagnostic numeracy assessment. Error bars show the 95% confidence intervals

Due to the relatively small sample sizes of 13 and 20 in the C2 NANN and C1 NA cohorts respectively, it was not possible to show that the additional fluency gains made by the students who used Numeracy Ninjas, as shown in Figure 8, were statistically significant. The gains in key skills by the C2 NANN cohort were close to passing the significance test ($p = 0.07$). The $p$ values for the additional gains made by the C2 NANN cohort in the mental strategies and timestables sections were 0.94 and 0.33 respectively.

Whilst there is not sufficient evidence yet to verify the statistical significance of the impact Numeracy Ninjas has had on our KS2 level 3 students, it is clear the results are suggesting there likely is a positive additional impact and more data should be collected in the future to support further statistical significance testing.

Figure 9 shows the fluency gains made by the KS2 level 4 students in the C1 NA and C2 NANN cohorts.
KS2 level 4 students were the prior attainment group who gained the most from Numeracy Ninjas. In comparison with the fluency gains made by the KS2 level 4 C1 NA control cohort, the gains of the KS2 level 4 students in the C2 NANN cohort in mental strategies, timetables and key skills were all statistically significant with the following hypotheses passing a two-tailed student t test with $\alpha = 0.05$:

1. Numeracy Ninjas increased the average KS2 level 4 student’s result in the mental strategies section of the diagnostic numeracy assessment by at least an additional 69% when compared to the gains made by the C1 NA control cohort.

2. Numeracy Ninjas increased the average KS2 level 4 student’s result in the timetables section of the diagnostic numeracy assessment by at least an additional 470% when compared to the gains made by the C1 NA control cohort.

3. Numeracy Ninjas increased the average KS2 level 4 student’s result in the key skills section of the diagnostic numeracy assessment by at least an additional 140% when compared to the gains made by the C1 NA control cohort.

Critics to comparing the KS2 level 4 students’ results in progress terms would argue that the level 4 students in the C2 NANN cohort were weaker than the level 4 students in the C1 NA cohort in all three areas of the diagnostic numeracy assessment at the start of the academic year. Therefore, it could be suggested that it was unfair to consider the proportion of the progress the C2 NANN level 4 students made on the diagnostic numeracy assessment over the year that raised them up to the proficiency that the C1 NA level 4 students had at the start of the year. I.e. perhaps this proportion of the progress that the C2 NANN level 4 students made would have been made without Numeracy Ninjas? Whilst we believe this to be unlikely, if this line of argument is followed it would therefore also be appropriate to
compare the results of the KS2 level 4 students also in absolute attainment terms, rather than just by progress. By inspecting Figure 9- Key Stage 2 level 4 students' mean scores on the diagnostic numeracy assessment. Error bars show the 95% confidence intervals it is clear that despite the C2 NANN level 4 students starting from a weaker position, they ended the academic year with a higher attainment in all three areas of the diagnostic numeracy assessment than the C1 NA level 4 students.

Even if the cautious view about differences in starting points is taken, it was found that the overall attainment of the KS2 level 4 students in the C2 NANN cohort was statistically significantly higher than the KS2 level 4 students in the C1 NA cohort in all three areas of the diagnostic numeracy assessment at the end of the academic year.

**These results strongly suggest that Numeracy Ninjas can have a statistically significant, transformative impact on the numeracy fluency (and the fluency of pre-GCSE Mathematics skills) of KS2 level 4 students.**

Figure 10 shows the fluency gains made by the KS2 level 5 students in the C1 NA and C2 NANN cohorts.

![Figure 10](image)

**Figure 10-** Key Stage 2 level 5 students' mean scores on the diagnostic numeracy assessment. Error bars show the 95% confidence intervals

Fluency gains by the KS2 level 5 students using Numeracy Ninjas were not statistically significantly different to the KS2 level 5 students in the C1 NA control cohort. We believe these results may suggest:
1. There could be diminishing returns gained by KS2 level 5 students in terms of the mental strategies and timestables; they are already reasonably proficient in these areas and therefore have less to gain.

2. If Numeracy Ninjas is not used, these students appear to make stronger gains than KS2 level 4 (and below) students in terms of their mental strategies, timestables and key skills. These students seem not to require Numeracy Ninjas to support the fluency gains to such an extent as KS2 level 4 students; they may acquire similar gains by following conventional lessons without the need for the additional retrieval opportunities given by Numeracy Ninjas.

As the KS2 level 5 students in the C2 NANN cohort likely performed better than the C1 NA level 5 students at the start of the academic year (see the 95% confidence intervals) the previous argument to compare attainment rather than progress (as per with the KS2 level 4 students) would not be valid and so was not analysed. A critic would argue, they started higher so they ended higher. Nonetheless, the C2 NANN level 5 students did end the year with higher attainment in all three areas. Further experiments are needed before we can reasonably claim Numeracy Ninjas leads to higher attainment on the diagnostic numeracy assessment for KS2 level 5 students.

However, we believe that KS2 level 5 students may still benefit from participating in Numeracy Ninjas; the results are inconclusive, but the question remains open. In terms of the impact on their average fluency gains across mental strategies, timestables and key skills, these may be modest, but an occasional use of Numeracy Ninjas could support the identification and subsequent closing of any specific basic numeracy skill gaps that KS2 level 5 students have. Nonetheless, we believe the frequency with which Numeracy Ninjas is used with these students should be considerably much less than the KS2 level 4 (and below) students.

Figure 11 shows the fluency gains made by the KS2 level 6 students in the C1 NA cohort.
There were no students in the C2 NANN cohort identified as KS2 level 6 and therefore the impact of Numeracy Ninjas on these students cannot be assessed. However, it is clear from the C1 NA control cohort’s results shown in Figure 11 that the conclusions for the KS2 level 5 students are also reasonable for KS2 level 6 students; there is only a small margin of possible improvement in the mental strategies and timestables which they are making from following their normal curriculum. They are also making substantial fluency gains on the key skills without using Numeracy Ninjas.

Ninja scores over time

Figure 12 and Figure 13 show the weekly mean Ninja Score by KS2 prior attainment group for the C2 NANN and C1 NA cohorts respectively.
The Ninja Scores in both cohorts improved across all prior attainment groups up until approximately week 11 when they start to plateau, broadly remaining constant for the remainder of the year. However, in both cohorts the Ninja Scores of the KS2 level 3 students (and the KS2 level 4 students in the C2 NANN cohort) rise again after the initial plateau. This suggests that the plateau was not caused by a definitive learning block; these students still had the capacity to improve.

The plateau and subsequent improvement raises a number of interesting questions about the most effective way to run Numeracy Ninjas in the classroom. What happened around week 11 (Christmas time) that led to a plateau in improvement for all prior attainment groups? We hypothesise that a possible explanation could be that the available benefits of *the testing*
effect had been reaped and that with the remaining topics that students were getting incorrect, the students were not going to improve just by repeated testing and seeing correct answers; they needed teacher input to learn how to correctly answer their outstanding topics. Some teachers did come to this conclusion themselves and took action with their classes to identify which topics students continually struggled with and acted to talk students through examples. However, this was not consistent across the department’s practice. It could, nonetheless, suggest a possible explanation for the subsequent progress of the aforementioned prior attainment groups after their initial plateau.

We believe that monitoring the C2 NANN and C1 NA cohorts’ Ninja Scores has been valuable as it suggests a format for improving our implementation of Numeracy Ninjas in subsequent years. After an initial approximate 10-week period, the focus should switch to diagnostic use to identify topics that remain outstanding and teaching tailored to specifically address these.
Concluding remarks

We believe that taking such a rigorous approach to analysing the impact Numeracy Ninjas has had on our Key Stage 3 students’ numeracy fluency at Wyvern College has given us confidence that it having a highly beneficial effect on the mathematics progress of our KS2 level 4 students (and most likely our KS2 level 3 students too).

More data collection is needed before we can say with confidence that it impacts positively for our KS2 level 2 students.

We will also adjust the implementation of Numeracy Ninjas next year in a number of ways:

1. All KS2 level 4 (and below) students will use Numeracy Ninjas every lesson until they regularly achieve Ninja Scores of 30.
2. In approximately week 11 we will ask teachers to consciously focus on using it diagnostically to identify outstanding gaps students have and to teach them how to approach these questions. As ever, a tool is only ever effective if there is a highly skilled craftsperson using it... In this context an expert mathematics teacher using high quality assessment for learning.
3. Level 5 and 6 students may use Numeracy Ninjas on a weekly basis from the start of the year with teachers being proactive about diagnosing and fixing any numeracy gaps from day one. Once they are satisfied these have been secured they will stop using Numeracy Ninjas and use other spaced-learning based starters with a higher pitch. We believe that the plateau in the Ninja Scores, if it was attributable to reaching a limit with the testing effect by students not being able to infer correct approaches from answers without explicit instruction, may suggest that any higher-pitch spaced-learning activities used subsequently should be based on topics previously studied in class, rather than unstudied topics.

Whilst these conclusions have been reached based on our own experiences and evaluation of running Numeracy Ninjas in our own department this year, we are fully aware that due to the flexible nature of Numeracy Ninjas, that other departments may have had success by using it in other formats. We fully encourage the sharing of your experiences, preferably backed up by data such that you have confidence in any suggestions you put forward. If the Numeracy Ninjas community can learn from each other all our students will be better off for it.

As ever, the journey is just beginning and we put forward the following suggestions, questions and challenges that the Numeracy Ninjas community may like to address collaboratively in the future:
1. What if the diagnostic testing were adapted so that it was time-based rather than binary ‘correct/incorrect’? If you created a diagnostic test that recorded the time it took the student to answer then compare it against whole cohort, national or international average times might it give us a better insight? Would we interpret the evaluation data differently?

2. Would developing some instructional resources, for example a series of video lessons, that students could access to learn how to approach questions they have not yet mastered on Numeracy Ninjas avoid any plateau in Ninja Scores (before 30!)?

3. Could students collaborate with class peers to learn how to approach questions they have not yet mastered? Could you identify and use ‘Ninja experts’ to help you support other students?

4. Is it possible to design a robust study to assess the impact that Numeracy Ninjas has on students’ GCSE Mathematics outcomes? Are there too many confounding variables? What about looking at the impact on progress and attainment in just Year 7 or 8?

5. How could we measure the ‘transferability’ of the improvement in skills we measured in the diagnostic numeracy assessment? Can students apply their fluency with these skills in different contexts? How can we tell?

6. Could you collect enough reliable data to test if Numeracy Ninjas’ impact on the numeracy fluency of KS2 level 2 students is statistically significant?
Bibliography


Department for Education. (2010). *The National Strategies- Primary- Teaching Children to Calculate Mentally*. Department for Education.


