

DoNow+

Greenshaw High School

Problem: What challenges do your school(s) have that need to be addressed?

The demands of the new GCSE and A-level qualifications put a much greater emphasis on retaining considerable content knowledge in the long term memory. Many schools have focused on developing different ways to address this increased rigour, changing their curriculum plans and sequences, as well as altering their teaching methods. A lot of these changes take significant time and resources and there is limited evidence as of yet to their impact on student learning. We wanted to develop an innovation that exploited the boost in student learning through regular retrieval and elaboration without having to wait years to rework our curriculum.

Innovation: How will the innovation help improve the problem you have identified and benefit teachers and learners?

Every lesson will begin with 10 short answer recall questions designed to test students' substantive knowledge from different parts of the curriculum. The questions are organised thus:

- Q1-2 from last lesson or recent learning
- Q3-4 from last week's learning
- Q5-6 from the last topic or unit
- Q7-8 from last term
- Q9-10 from last year

The first two questions are primarily designed to support responsive teaching. They relate to the learning which is most recent and upon which the next steps of learning (today's lesson) will be based. William's work on formative assessment highlights the importance of ensuring we check students' understanding of prior learning (Q1&2) before we proceed (William, 2014). DoNow+ helps teachers to complete this check in a systematic but time efficient manner. While students are tackling the ten quiz questions, the teacher circulates and checks students' responses to just the first two questions and on this basis, determines the next steps in the lesson:

- **If 80% or more of students** get the first two questions correct then the teacher can be confident that they are ready to move on. At this stage, the teacher will ask a question which requires students to elaborate on the understanding encompassed in those first two questions because McDaniel and Donnelly (1996) have shown that elaboration (i.e. meaning-enhancing additions, constructions, or generations) has a powerful enhancing effect on memory
- **If less than 80% of students** get the first two questions correct then the teacher makes a further choice:
 - a) To address the misunderstanding **immediately** through re-explaining, using examples or other students to help explain. This clarification exercise should be pre-planned based on the teacher's knowledge of common misconceptions and therefore likely errors and roots thereof.
 - b) If the misunderstanding is too deep/complex/unclear then **postpone** addressing it to the next lesson.

Questions 3-10 are more focused on the development of memory. The regular testing embodied in this innovation is informed by Roediger and Karpicke (2006)'s work on the power of testing to boost memory retention. By linking the questions to different parts of the curriculum, e.g. Q4 is from last week while Q8 is from last term, we are applying Baddeley (2012)'s findings on the positive impact of distributed practice and Bjork and Bjork (2003)'s research on desirable difficulty for boosting memory retention. The decision to focus on building up the levels of background knowledge in coherent networks is informed by the work of Hirsch (1988), Sweller (1994) and Willingham (2002). Given this, our schools have already defined 'knowledge organisers' which list the substantive knowledge related to each topic in the curriculum and it is this knowledge which the DoNow+ questions will interrogate. Students will NOT have access to the knowledge organisers during the quiz.

The delivery of the quiz is also important. While Q1&2 are checked (and responded to) by the teacher, Q3-10 are self-checked by students (the answers are displayed on the board) which makes DoNow+ more time efficient but also leverages the research on the impact of hypercorrection and metacognition on enhancing learning.

An extension question will be included for those who may finish the first 10 questions early thus ensuring no learning time is lost. The entire DoNow+ should be completed in no more than 10 minutes in order to leave sufficient time for new learning in the lesson. Although this is a demanding time schedule, we have trialled this procedure in one class this year and have proven that it is doable. The short, sharp design of the questions facilitates speed of completion as does the fact that likely misconceptions and follow up questions have been premeditated by the

teacher. Having said this, if on some occasions the quiz is taking too long, it will not significantly undermine the process if the teacher decides to end it before all students have completed all the questions and indeed this situation is not uncommon in a mixed ability teaching context.

Existing evidence: What evidence is there that this innovation will improve outcomes?

There is a wealth of existing evidence from the field of cognitive science about the benefits of testing, spacing and interleaving in developing long term memories and helping to form coherent schema (Dunlosky et al, 2013). The work of Bjork and Bjork (2003) on Desirable Difficulties, for instance, as well as the likes of Roediger and Karpicke (2006) on the testing effect, have been hugely influential in showing ways to slow down short term performance in order to increase long term retention and understanding. Similarly, Sweller's work (1994) on cognitive load has demonstrated the importance of drawing from well-formed memory traces in organised neural networks to reduce demands on working memory when engaged in more complex, multistep cognitive tasks.

William's extensive work on formative assessment, particularly through the use of targeted diagnostic questions, has shown the benefits to learning of teachers anticipating misunderstandings and using their students' responses to guide their next steps (e.g. William, 2014).

Research question or hypothesis: What effect will the intervention, implemented for how long, with which pupils, have on what outcomes?

How does the use of ten highly structured recall questions and one carefully planned elaborative question at the start of every lesson, over two terms, improve the performance of Year 8 students when responding to analytical questions in History and in English, compared to pupils who do not have these tests?

Method: Include sample, design, measures, intervention, process evaluation, and analysis

Sample/ participants

Two Year 8 classes with a broad range of prior attainment (although not including pupils with the lowest prior attainment) in each of History and English in each of the two schools will form the sample.

Although a Do Now is already common practice in both schools it is very different innovation we propose in terms of the structured nature of the questions, the spacing and interleaving function, the review process, the pre-planning of likely misconceptions and necessary responses, etc.

None of the teachers or students in either school have ever experienced anything like this before.

Since we will be collecting and analysing student data, we will need parental consent. This will be gathered using an Opt Out process. We will use an adapted version of the letter template provided by the IEE.

Design and assignment to condition

We require a control and intervention group in each school because the two schools will be teaching different curriculum content and therefore measures of impact must be made within each school. We will use the baseline test and teacher characteristics to ensure the two classes are similar. The baseline tests will use Comparative Judgement to measure students' performance in responding to analytical questions. We will average the results of two baseline tests to create as fair a baseline as possible. Within each pairing of classes, we will then randomly select one class to be the control and one class to be the intervention.

Students are in different classes for History and English and therefore it is possible that they might be in the intervention class for one and the control for the other but this should not matter since our premise is based on the idea that by increasing a student's knowledge in a particular subject we can improve their attainment in that subject. The focus on subject knowledge leaves little to no room for cross fertilisation across subjects which would not be the case if the focus was on skills.

To help ensure that the teachers in the intervention and control groups are similar we will define, in advance, a set of criteria against which to measure them, thus making this process as objective as possible. Criteria will include: Number of years teaching, whether or not the teacher holds a middle or senior leader position, their past experience of using low stakes quizzing in a systematic way. No teacher could teach both intervention and control groups due to the staffing structure in both schools.

Measures

We will measure the impact of the intervention by asking students to respond, in writing, to an analytical question typical of the sort they normally face in History and English. The analytical questions will be set, in advance, by an independent source, likely a colleague who is not involved in teaching the Y8 course. The questions will be kept confidential from the teachers involved in the evaluation to remove the chance of teachers 'teaching to the test'.

Students will complete these assessments twice before the intervention, once during and once after. The interim (during) test will be solely for the purposes of process evaluation while the after (post) test will be used to judge our impact. The questions will be different in each assessment because the subject matter being studied at different points in the year will be different but the nature of the question and the demands it makes on the student will be similar each time.

The responses of all students, intervention and not, will be evaluated using Comparative Judgement. CJ ranks students relative to the performance of others who have completed the same task and thus it fits our purpose well. Functionally, CJ will also allow us to utilise the subject teams across both schools to make the judgements on each student's work. To elaborate: when it is time to mark the Greenshaw History assessments, CJ allows us to make these available to whomever we choose so we plan to gather the teaching teams from both schools together and have them all judging the Greenshaw scripts. This should increase objectivity as some of the judges will have absolutely no connection whatsoever with the answer they are judging. Even though some teachers might recognise a particular student's handwriting, any consequent bias in their judgement will be watered down due to the number of other judgements, on the same piece of work, made by people who do not know that student. Working together on the judgement process will also increase the number of judges which reduces the number of judgements each judge must make and this decreases the time taken to mark each assessment (although this time will be spent later by both teaching teams marking the assessments of the other school). Finally, it gives the teachers in the other school the chance see the work of students the same age as their own students but in a different school and thus should support them in calibrating their broader judgements and expectations of their own students.

CJ will also give us a scaled score. If our premise is correct then those students in the intervention classes will achieve higher scores in the CJ process after the intervention than they did in the baseline assessments.

Intervention

In the autumn term, lessons will be conducted as normal but two assessments will be conducted

and averaged to establish the baseline performance of students in responding to analytical questions.

Following the baseline tests, classes will be assigned either to intervention or control and then the intervention teachers will attend a 1-day induction workshop with the project manager to explain the rationale and methodology of the innovation. They will be taught how to create and deliver the quizzes and will collectively create a bank of quizzes for the first few weeks of lessons to ensure that these quizzes meet the requirements of the innovation. This will give teachers a head start on the quiz creation but after that, they will have to create their own. A central repository will be set up (possibly utilising GoogleDocs or something similar) in which all questions will be stored. This will allow the project manager to regularly check that the quizzes are being constructed correctly. It will also allow teachers across the intervention groups to see each other's work which is a positive in terms of supporting consistency across the groups and also encouraging and supporting each other in terms of types of question.

Intervention teachers will return to their schools and will, for the next three half terms, begin every lesson with the DoNow⁺ intervention.

The project manager will visit lessons in week two of the intervention to check that the DoNow⁺ is being delivered with fidelity and again in week four and then once every half term. If issues are identified, the relevant teachers will receive further coaching from the project manager and be revisited to check that the issues have been resolved.

Intervention teachers will meet once every half term (0.5 days), to share their experiences in order to maximise fidelity and commonality of practice across the different intervention classes. These sessions will also be used to evaluate the process and identify any potential/real issues which might need to be addressed but it will be important to emphasise to all participants that the experiment must be conducted according to the terms defined in this proposal document for the duration of the whole experiment, otherwise its validity will be undermined. Finally, the sessions will be used to support teachers in the development of the next tranche of quizzes.

Process evaluation

The project manager will conduct observations on the intervention classes to check for fidelity in the design and delivery of the DoNow⁺, e.g. the style and structure of the questions and the teacher's response to the outcomes of Q1&2.

She will visit the intervention groups in weeks 2 and 4 and then every half term. She will visit the control groups in week 2 and then every half term.

Observations will follow a predefined schedule to ensure consistency of judgements across the

different observations.

In the final training session, intervention teachers will complete a questionnaire around the process which will include questions such as:

- What benefits do you perceive there to have been from this innovation?
- What costs do you perceive there to have been?
- What factors were supportive of this innovation?
- What factors served as barriers?
- Would you make any changes to the innovation?
- Will you continue to use this intervention now that the project is finished?

We will pilot this questionnaire with a teacher not involved in the project in order to ensure it makes sense to teachers and that we are asking the questions we think we are.

We will use SurveyMonkey for this questionnaire as we have experience of this tool already.

A questionnaire will also be used, at the end, to gauge the intervention students' feelings about the process of the DoNow+.

This questionnaire will be done in class (to make sure it gets completed) and will be paper based (due to the challenges of IT access for students) however we have the facility to have this 'marked' electronically which will speed up the analysis process.

Data analysis

We will calculate effect sizes using the following formula:

$$\frac{\text{Mean progress of intervention group} - \text{mean progress of control group}}{\text{Standard deviation of progress of whole sample}}$$

Separate effect sizes will be calculated to compare the progress of intervention and control group pupils in History and English at each school.

We will look at the data through the usual lenses applied in school, i.e. Pupil Premium, gender, ethnicity and prior attainment.

We will use thematic analysis to evaluate the process data.

**Conclusion: What will happen if your innovation improves outcomes, or not?
What are the limitations of your evaluation?**

Potential Limitations

It is difficult to isolate the impact of the innovation from other variables between classes but by using teacher characteristics and the baseline tests to select the groups we should minimise this.

Even if a control class experienced a small element of the intervention this would not be unduly damaging to the experiment given the complexity of the full DoNow+ package.

Implications

Positive: Extend the trial to other subjects and year groups.

Negative: We will review the design of the innovation, e.g. the number of questions, to see if this delivers more positive results. If this also fails, we will investigate alternative strategies to build in retrieval practice, elaboration and responsive teaching in a manageable and effective way.

References

Baddeley, A. (2012). Working Memory: Theories, Models, and Controversies. *Annual Review of Psychology*, 63, 1 - 29.

Bjork, E. L., & Bjork, R. A. (2003). Intentional forgetting can increase, not decrease, residual influences of to-be-forgotten information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29(4), 524-531.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving Students' Learning With Effective Learning Techniques: Promising Directions From Cognitive and Educational Psychology. *Psychological Science in the Public Interest*, 14(1), 4–58.

Hirsch, E.D. (1988). *Cultural Literacy: What Every American Needs to Know*. Random House USA Inc.

McDaniel, M. A., & Donnelly, C. M. (1996). Learning with analogy and elaborative interrogation. *Journal of Educational Psychology*, 88(3), 508-519.

Roediger III, H.L., and Karpicke, J.D., (2006). The Power of Testing Memory: Basic Research and Implications for Educational Practice. *Perspectives on Psychological Science*, 1, 3, 181 - 210.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, 4, 4, 295 - 312.

Willingham, D. (2002). *Inflexible Knowledge: The First Step to Expertise*. Available at <https://www.aft.org/periodical/american-educator/winter-2002/ask-cognitive-scientist>

William, D. (2014). *Formative assessment and contingency in the regulation of learning processes*. Presented at *Toward a Theory of Classroom Assessment as the Regulation of Learning* at the annual meeting of the American Educational Research Association, Philadelphia, PA, April 2014.

