

Evaluating the impact of feedback using an audio tool compared to written feedback

Notre Dame High School, Sheffield

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

Teachers spend a long time providing written feedback to pupils. We will evaluate the effectiveness of alternative ways of providing feedback.

There is limited evidence as to what type of feedback is most effective in moving pupils forward. Improvements in technology, development of apps such as *Kaizena* and widespread ownership of mobile devices means the recording of audio feedback is now a practical and viable alternative to providing written feedback. The effectiveness of this alternative form of feedback should be evaluated.

Teachers spend a significant proportion of their time marking work and workload is a major reason cited for current recruitment and retention issues (DfE, 2015). Evaluating the impact of alternative forms of feedback on teacher workload is therefore important for the entire profession.

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

We will evaluate the effectiveness of teachers recording verbal (audio) feedback for pupils. This will involve teachers recording verbal feedback digitally and sharing this with pupils via, for example, school email instead of providing more traditional written feedback.

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

Literature Review

“Despite its centrality to the work of schools and teachers, there is in fact little high-quality research related to marking.”

“There is an urgent need for more studies so that teachers have better information about the most effective marking approaches.”

A Marked Improvement? (Elliott, et al., 2016) This paper cites a study that found that university pupils often did not understand the terms used in written feedback. This resonates with our own anecdotal evidence where teachers felt that pupils were better able to understand audio feedback. There is also reference to wider evidence concluding that the specificity of the feedback has an impact on performance. Our experience is of teachers being able to give more detailed (ie more specific) feedback in a given period of time via verbal (audio) feedback.

The Power of Feedback (Hattie & Timperley, 2007) cites a meta-study that suggests the effect size of feedback is most when the feedback gives the learner cues and is video or audio in format. This meta-study was carried out in 1999 before easy and efficient access to mobile technology. Our research will investigate whether audio feedback given by a teacher **in the context of a busy working schedule** is more effective than written feedback. (ie do these findings translate to ‘less than perfect’ conditions, where teachers have limited time in which to provide feedback (written or audio)?)

Pilot

Teachers of sociology and mathematics at Notre Dame have trialled giving audio feedback to sixth form pupils with very positive responses. Access to mobile devices is sufficiently good that this alternative form of feedback was readily accessible to pupils. A pupil voice survey indicated that the majority of pupils scored the usefulness of the audio feedback higher than written feedback.

Teachers felt that pupils were better able to understand the audio feedback, found it easier to give more audio feedback in a given time, and were better able to pick up on verbal cues and tone of voice than from written punctuation. A pupil voice survey indicated that the majority of pupils said they valued being able to listen back to feedback from their own teacher.

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

The specific primary question we will address is:

Is verbal (using an audio tool) feedback, delivered over a discrete topic, more effective than written feedback in improving test outcomes in sociology and mathematics A-level?

And an important secondary question:

Does providing verbal (using an audio tool) feedback to pupils rather than written feedback have a positive impact on teacher workload?

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

Sample

Our trial will involve 32 groups of Year 12 pupils across eight schools and limited to two contrasting subjects: sociology and mathematics.

Eight partner schools will each nominate two maths and two sociology Year 12 groups to participate. We have chosen maths and sociology as these are amongst the most popular A-level subjects and schools are likely to have multiple groups of Year 12 pupils in these subjects. This provides a total cohort size of approximately 640 pupils.

Consent

We will inform parents of pupil participation in the trial.

As part of our training, we will ensure that all staff involved know to give pupils the opportunity to verbally opt out of having their work analysed as part of the comparative process.

Allocation to groups

One group per subject per school will be randomly identified as the control group (this group will receive written feedback) and the other will be the treated group (who will receive verbal feedback). This allocation of treatment has been chosen in order to minimise logistical challenges for participating schools.

The pupil groupings will be normal timetabled class groups. The allocation of class groups to control or treated cohorts will be carried out randomly by us. One maths class and one sociology class in each school will be randomly chosen as the treated group. In the unlikely event that pupil numbers are such that a school is not able to support two groups, then we will randomly assign the single group to either the control or treatment.

We will obtain prior attainment data that will allow us to ensure that overall, the control and treated cohorts are of approximately equal prior attainment. We will carry out a further random assignment if the overall prior attainments of the control and treated cohorts are not approximately equal.

Measures

For each school and subject (eg school "A" treated maths group and control maths group) we will rank the pupils before the trial (using prior attainment data) and again after the trial (using a comparative marking moderation process).

This will allow us to calculate a separate effect size for the treatment in each subject in each of the eight schools concerned.

The prior attainment data for maths will be uniform mark scores from GCSE maths (where they are known), with the middle uniform mark score for each GCSE grade being used where the uniform mark score is not known.

The prior attainment data for sociology will be the average GCSE points score of each pupil.

The final ranking after the trial will be based on a comparative marking of a final written assessment. The pupil work from the assessment will be copied, with one copy being used for normal feedback (to help minimise disruption to normal teaching practices) and the second copy being anonymised and saved for later moderations. After the 2018 exam season (to remove barriers for attendance), participating teachers will collectively moderate the pupils' final assessments via the No More Marking comparative judgement process.

The final moderation will be done blind, with teachers comparing anonymised control and treated work from schools other than their own.

We will also collect qualitative data from pupils via Likert scale surveys following each piece of marked work. The questions will specifically address the effectiveness of the feedback to the pupils and reference known characteristics of good feedback as outlined in *What Makes Great*

Teaching (Coe, et al., 2014)

During the trial, all pupils will complete a Likert scale questionnaire to obtain qualitative data on the effectiveness of the feedback. These questions are motivated by Hattie & Timperley (2007).

Qualitative questionnaire for pupils during the trial:

- I understand the feedback that I've been given
- The feedback helped clear up any misunderstandings of what the question was asking me to do. (FT)
- The feedback helped by giving me specific things I could do to improve the work. (FP)
- The feedback told me the areas I could make improvements in, but left me to work out exactly what those improvements should be (FR)
- Overall this feedback was useful

The questionnaire will also have an open question allowing for free text responses.

The initials in brackets refer to the three 'levels' of feedback that Hattie's *The Power of Feedback* argues are the most powerful.

FT=feedback aimed at the task

FP=Feedback aimed at the process

FR=feedback aimed at self-regulating

Innovation

Subject specialists will identify a maths topic and a sociology topic that will be taught by all exam boards at some point during Year 12. The topics will be large enough to allow 2–3 pieces of work to receive either control feedback (written) or treated feedback (recorded verbal).

- Preliminary input from subject specialists points towards a mathematics topic covering *index laws, evaluating powers and differentiation*. This is a suitably-sized topic, appearing on all AS specifications. It is also a topic where pupils who are not able to respond effectively to feedback will make poor progress that will be reflected in the final summative assessment.
- For sociology, subject specialists have suggested *Research methods* as an appropriate topic to use for the trial. We will restrict the trial to centres studying AQA sociology (most

commonly taken specification), as the variations between case studies used by pupils studying different sociology specifications may add complexity to any final moderation process. Although there may be a slight variation in the depth of study of this topic (as those centres entering all pupils for AS may cover this in slightly less depth in Year 12 than other schools), these variations will be present in both the control and treated groups for any schools concerned.

Training

Training for all participating teachers will be carried out after the 2017 summer exam period to remove barriers to teacher attendance.

Pre-trial training schedule:

- Parameters for the trial
- Administration of questionnaires during the trial
- Running the final assessment
- Dealing with consent and confidentiality
- Dealing with the technical aspects of giving audio feedback
- What to do if there any problems during the trial
- An opportunity for clarification on any potential issues

Innovation

Participating teachers will be asked to provide 'appropriate feedback' on the work they are marking. This will avoid potential issues around whether they should set out to either spend the same time giving audio feedback as they would written feedback, or see whether they can give the same amount of feedback more quickly via an audio recording than via written comments.

Participating schools will also agree that dedicated improvement and reflection time will be built into lessons so that pupils are able to respond to the feedback. Participating schools will agree that, apart from the way in which feedback is given, the two groups for each subject will otherwise follow as similar a scheme of work for the topic concerned as possible.

Process evaluation

We will collect qualitative data from teachers involved on and their perception of the workload

associated with it using a Likert scale questionnaire.

On completion of providing feedback for work in the trial units, all teachers will complete a Likert scale questionnaire to provide data on their perception of the impact of giving audio feedback on their workload. Although workload can be quantified in terms of actual time spent, it would be difficult to evaluate this accurately and we will therefore evaluate teachers' perceptions of the workload via the following question after each piece of marked work:

- What is your perception of the workload involved in providing feedback for this particular piece of work compared to the normal workload in providing feedback on work?

Data analysis

We will calculate effect sizes for the treatment for each participating school and for each subject. This will provide up to 16 separate effect sizes. We will work with IEE to draw overall conclusions from the larger dataset containing all participating schools.

We will analyse the difference between pupil and teacher responses to questionnaires after each piece of marked work in order to evaluate any 'novelty effect' that may be affecting the reliability of the qualitative data.

Confidentiality

Audio files will be recorded and shared on existing school communications platforms (eg school email).

Any shared pupil work will be anonymised.

The final report will anonymise schools and individual pupils as far as publishing data concerning the outcomes of the trial is concerned.

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

Implications of positive or negative outcomes

If the outcomes are positive we would look to broaden the study, to include different age groups and subjects, and also vary the length of time over which the treatment was given.

If the outcomes are negative, we would look carefully at the qualitative data collected to establish whether there was anything about the way in which the pilot study was conducted that had

prevented the pilot running as intended or limited the impact of the treatment.

Possible limitations and ways these are being addressed

To mitigate against resentment or demoralisation of those in the control group, we will emphasise during training that:

- The trial only works because there is a control group – they are equally as important as the treated group.
- The purpose of the trial is to evaluate whether or not the treatment is more effective. We don't know for certain whether it is or not.

To mitigate against the 'novelty effect' for the intervention group, we will compare the qualitative pupils' responses after each of the three pieces of work receiving feedback. If there is a novelty effect, this would be seen in pupils' responses being less positive after the second and third pieces of work.

We are managing the risk of biased marking by ensuring no teacher makes any comparative judgements involving their own pupils.

The risk associated with diffusion of treatment are minimised as the two treatments are sufficiently distinct. Therefore there is minimal risk even if a single teacher within the same school is teaching both the control and treated group.

We are ensuring that our treated and control cohorts are overall approximately equal in terms of prior ability to mitigate against external factors skewing the results.

Communication

Eight schools involved from two geographic regions, mean that results will be quickly disseminated across a large number of schools.

We will liaise with HRS and Learn Sheffield to identify opportunities to share the findings of the pilot study.

References

Coe R, Aloisi C, Higgins C & Elliot Major L, 2014. *What Makes Great Teaching*, England: Sutton Trust.

DfE, 2015. *Government Response to the Workload Challenge*, London: Department for Education.

Elliott V et al., 2016. *A Marked Improvement*, England: EEF, Oxford University.

Hattie J & Timperley H, 2007. The Power of Feedback. *Review of Educational Research*, 77(1), pp. 81–112.