



Transforming Assessment

Using Adaptive Comparative Judgement as a reliable way to assess oracy at scale

Case Study

Voice 21

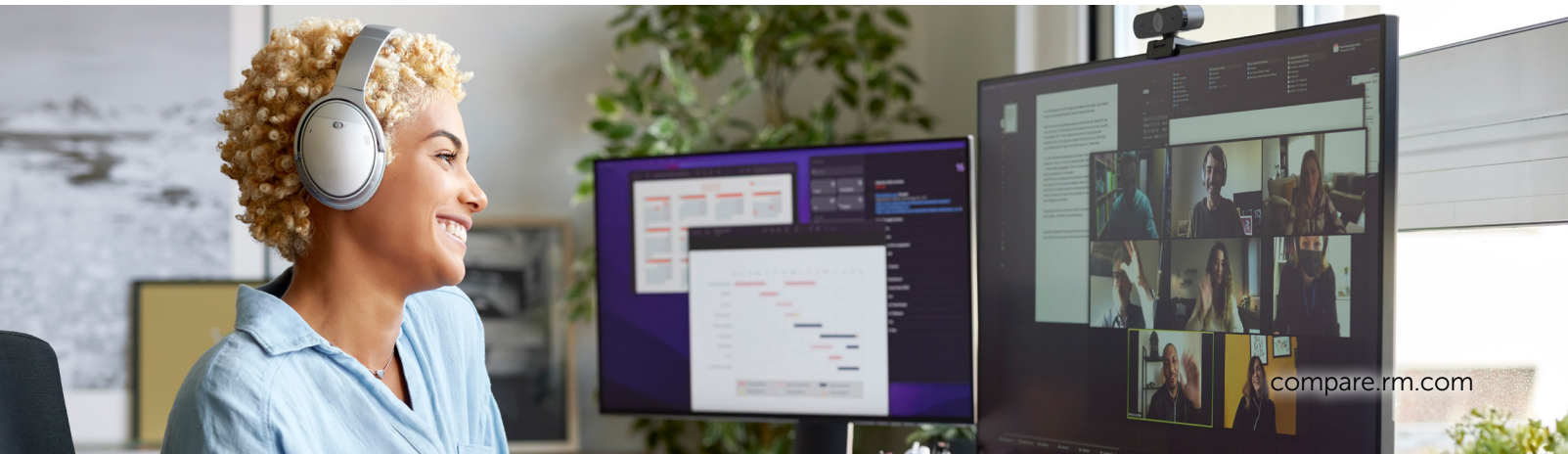
Introduction

In schools all around the world, oracy is a powerful tool for learning. By teaching students to become more effective speakers and listeners, we empower them to better understand themselves, each other, and the world around them. It is also a route to social mobility: empowering all students to find their voice to succeed in school and life. Voice 21, the UK's oracy education charity, work with teachers and schools around the country to improve access to quality oracy education, with the aim of creating a more equitable space where every voice is valued and heard.

Despite wide recognition of its importance, oracy has struggled to get the attention it deserves, and its assessment has been challenging. Assessing the spoken word is logistically challenging, and the authenticity of existing assessment techniques raise concerns over validity, while high levels of subjectivity means that reliability is hard to achieve. A key understanding is that oracy is much closer to a performance and is therefore

better suited to holistic assessment techniques. Voice 21 partnered with Assessment from RM to explore the use of Adaptive Comparative Judgement (ACJ) to assess oracy at scale and in a more efficient and reliable way.

Comparative judgement allows assessors to compare two pieces of work side-by-side and decide which is better against a holistic statement. This is done multiple times by multiple assessors to ensure a high level of reliability. RM Compare is a digital tool that makes this process more efficient, adapting the algorithm in real-time to decide how often each piece of work needs to be seen and judged based on the results of previous rounds of comparison.



The Challenge

The All Party Parliamentary Group (APPG) for Oracy 2021 Report '[Speak for Change](#)' identified 5 reasons why oracy matters:

1. Improves academic outcomes.
2. Underpins literacy and vocabulary acquisition.
3. Supports wellbeing and confidence.
4. Enables young people to have access to employment and thrive in life beyond school.
5. Develops citizenship and agency.

The APPG report recognised that while oracy education matters for everyone, it is disproportionately advantageous to groups including those experiencing poverty, children with speech, language and communication needs (SLCN), and those with special education needs and disabilities (SEND).

The place of oracy education in the school curriculum has been marginalised because of the challenges presented by current methods of assessment. The underlying skills and competencies of oracy are best assessed through a holistic approach, instead of absolute judgements. Although some effort has been made to take a holistic stance using a standards framework, the challenge of reliability remains. Without a reliable oracy assessment, teachers cannot take an evidence-based approach to the teaching of oracy; and school leaders struggle to make a case for oracy to receive the time or resource it requires.



The solution

Comparative Judgement allows oracy to be assessed holistically, respecting its performance qualities, subjectivity, and variability. RM Compare was used as an ACJ system to explore what this would enable schools and teachers to learn about students' oracy.

The tool needs to be reliable for schools to monitor their students' progress over time and compare their aggregate performance with other schools. In the longer term, this approach can be used to generate standardised rank orders of representative samples of students across the UK, which can then become the benchmark for future assessment. This would be a significant step towards ensuring that there is a comparable level of insight into schools' performance in the teaching and learning of oracy as for other important aspects of student attainment, such as literacy or numeracy.

The implementation process

A robust system

RM Compare is a cloud native product. It is built on Amazon Web Services (AWS) and has achieved the Partner Qualified Standard. To do this, it had to pass the AWS Foundation Technical Review (FTR) that requires providers to carry out a periodic review every few years of the system architecture and operations to identify gaps and continually improve. Specifically, it provides guidelines to adopt a subset of best practices to reduce risks around security, reliability, and operational excellence, as defined by the AWS Well-Architected Framework (WAF). RM Compare is built on a 'one-to-many' principle. It is designed to scale globally, making it applicable to different customers, organisations and environments.



Speaking the same language

A key requirement here was to focus on the practitioner's tacit knowledge. Specifically, the need for everyday classroom teachers to have confidence in their own daily judgement, sometimes called knowing 'what a good one looks like'. To do this, they would need to have a clear knowledge and understanding of effective oracy. RM Compare's use of comparative judgement met this need. It offered the possibility of establishing reliable oracy standards in a way which was authentic and valid, and had the potential to make the standard available for 'when-ready' assessment.

It was important that the technology allowed for the use of both video and audio content from several different sources. It is recognised that there are high levels of uncertainty here and a system that encourages experimentation and discovery is needed. RM Compare offered the technology that enabled this.

The logistics

Two sessions were completed with teachers from across the UK who were asked to judge students completing oracy tasks in two age groups. The software would surface a pair of students and the teacher would be asked to judge which one was 'best' against a holistic statement. The combined efforts of the judging pool produced a standardised rank order representing speakers from across the UK, allowing for a more reliable assessment of students' oracy than other available methods.

The iterative nature of the comparative process itself can also help teachers to develop their confidence and tacit knowledge of the oracy standard. By building teachers' accessibility and familiarity with oracy work from a national sample of schools, assessing oracy is made simpler and more reliable.



Authenticity as invaluable

Oracy tends to have a performance element to it which can be lost easily, and previous attempts to develop an oracy assessment have encountered significant challenges when released into a school context. Some of these challenges are practical such as how feasible the assessment is in a busy, noisy classroom, and others concern the validity of the assessment so for example, around perceived authenticity of the audience.

Therefore, it was important to use real students as subject items and real teachers in the judging process. While this was a challenge, it was essential in the validation process, as we had to be able to observe the way in which real students and teachers would interact with the assessment tasks and the comparative judgement platform.

The project expected to expose concerns regarding the use of oracy video images in school assessment. The responsibility of video creation was placed with the school participants to decrease the pressure on students by filming and ensure the overall authenticity of the items being assessed wasn't compromised by the digital interface. Although the quality of the items varied, and the participants in this project were early adopters, this proved invaluable in reviewing the project, providing key insights, and planning for the next. For instance, the key benefits of ACJ – efficient, predictable, and a higher

output and throughput – were validated to the extent that confidence to move to a subsequent project was achieved. In addition, the learning achieved informed the likely requirements needed for an on-demand solution.

Maintaining ethical practice

The use of video content, specifically, video content of children speaking, raises several concerns. Understanding and mitigating this was a key consideration for the project.

An early-stage project like this attracts early adopter participants who by their very nature, tend to have supportive approach behaviours. Even so, it was important for all participants to understand clearly the scope of the project and their role within it. The Data Controller in this scenario was Voice 21 who were able to select and manage a subset of their users, imposing necessary controls and safeguards along the way. As the Data Processor, RM Compare employs the principles of data protection by design and default. A key point here is that this is ongoing and is always a fundamental part of the learning and decision making.



The future

“Oracy can be assessed – Comparative Judgment - which relies on assessors making quick comparisons between videos of student talk- is a reliable way to assess oracy.”

Voice 21, Annual Impact Report

RM Compare is developed using agile principles and takes an interactive approach. As part of this project, the design team at RM spent over 50 hours with participants to better understand the user experience. The learning gathered was used to make multiple changes to the production environment, but to also produce several proofs of concept to further test hypotheses and assumptions.

A second, larger project is imminent, building on the learnings to date. This will provide further validation to

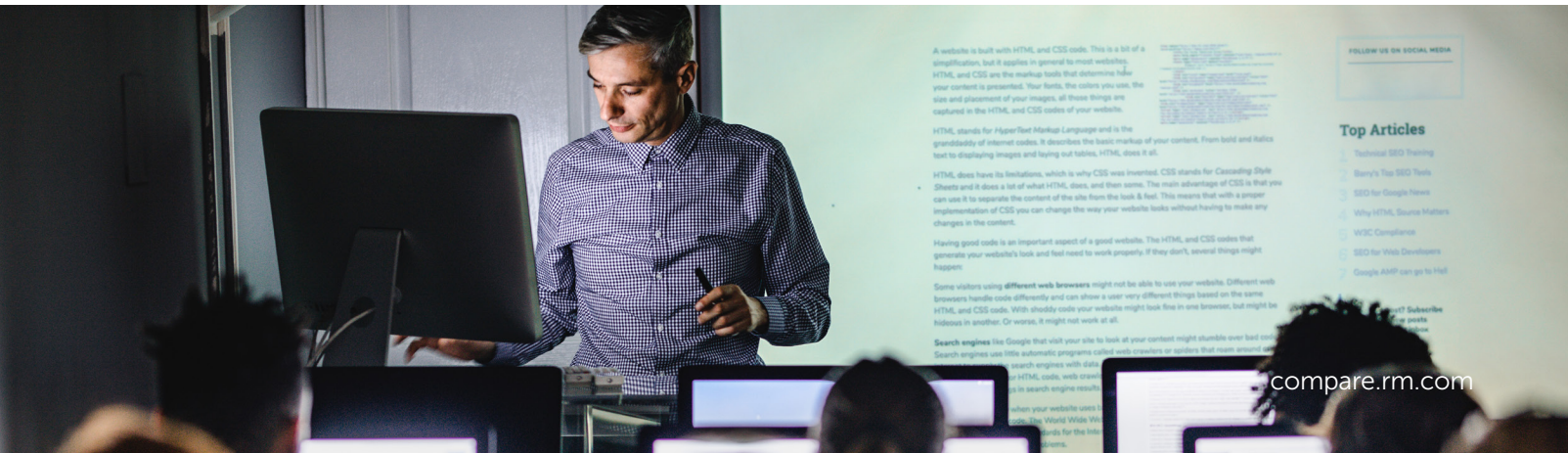
this case and uniform some of RM Compare’s long-term ambitions, specifically the development of the world’s first on-demand comparative judgement system. In doing so, we will be able to provide the thing that classroom teachers want more than anything – a ‘when-ready’ oracy capability.

The work undertaken has encouraged the RM Compare team to support other spoken work projects around the world, including the development of higher order presentation skills and language assessment more widely.

References

[Insights and Impact report 2021-22 - Voice 21](#)

[Oracy APPG - Voice 21](#)





**Interested in
finding out more?**

Learn more about digital assessment at
compare.rm.com