Interactive oral assessment for statistical literacy

Stephanie Budgett and Leila Boyle
The University of Auckland, New Zealand
s.budgett@auckland.ac.nz

In the quest to distinguish fact from fiction in a world awash with misinformation and disinformation, statistical literacy is more crucial than ever. One of the key characteristics of a statistically literate citizen is their ability to communicate and justify their reaction to data-based information encountered in various contexts. An interactive oral assessment (IOA) can be a valuable and authentic way of assessing these skills. IOAs involve genuine, unscripted conversations, facilitating opportunities for an instructor to query responses to evaluate students' understanding more fully and allowing students flexibility in demonstrating their knowledge. IOAs are also less prone to academic misconduct than their more traditional written or online counterparts. This paper presents the rationale for introducing an IOA in an introductory statistical literacy course, outlines the implementation process, and shares reflections from course instructors alongside preliminary student feedback.

INTRODUCTION

The following statement is from an opinion piece published in the *New York Times* (Harari, Harris, & Raskin, 2023):

> In 2022, over 700 top academics and researchers behind the leading artificial intelligence companies were asked in a survey about future AI risk. Half of those surveyed stated that there was a 10 per cent or greater chance of human extinction (or similarly permanent and severe disempowerment) from future AI systems.

Figure 1. Excerpt from the *New York Times*.

The data quoted arose from The 2022 Expert Survey on Progress in AI (see https://aiimpacts.org/2022-expert-survey-on-progress-in-ai/ for the survey report), which sparked significant media attention at the time (e.g., Klein, 2023; Lanier, 2023; Piper, 2022; Wallace-Wells, 2023). If the excerpt reflects the beliefs of 'over 700 top academics and researchers behind the leading artificial intelligence companies', how concerned should we be? To answer this, an inquiring and statistically literate reader could consult the survey report and fact-check the statement to discover that:

- The organisation responsible for implementing the survey contacted over 4,200 people who had published work at two large machine-learning conferences in 2021
- The response rate was 17%
- Each of the 738 (17%) respondents was given a random subset of questions, one of which was:
  What probability do you put on human inability to control future advanced AI systems causing human extinction or similarly permanent and severe disempowerment of the human species?
- This question elicited 162 estimates, the median of which was 10% or more

These facts generate more questions than answers. For example, we might ask, 'Who is missing?' (Harford, 2021), wonder about the basis for selecting these two 2021 conferences and how representative delegates who published work at these conferences are among the broader population of top academics and researchers behind leading artificial intelligence (AI) companies. Further, why was the response rate so low, and in what ways could those who responded to the survey differ from those who chose not to answer? Was the survey designed to force a response to this question, or did some who were given this question decide not to respond? How variable were the responses to this question? How confident can we be about the beliefs of AI academics and researchers based on the estimates provided by 162 people sampled as described? How would *you* interpret and answer this question? What would

In: Birk, L., Loth, G., Jotzo, L., Binder, K., & Frischemeier, D. (Eds.) (2026). *Statistics and Data Science Education in STEAM. Proceedings of the Satellite Conference of the International Association for Statistical Education (IASE), September/October 2025, Münster, Germany*.

you base your probability estimate on? A systematic and rational reflection of such questions will likely result in acknowledging the potential biases present in surveys like the 2022 Expert Survey on Progress in AI, leading to a more reasoned judgment on the statements made in the opinion piece.

Fostering students' statistical literacy is crucial in a world increasingly impacted by misinformation and disinformation. Statistical literacy, which relies on both knowledge and dispositional elements (Gal, 2002), enables critical engagement with news media articles such as the example shown in Figure 1. Although Engel and Ridgway (2023) note that "[m]any courses are not designed to teach relevant skills and improve learners' statistical literacy" (p. 17), there are notable instances of courses that have been purposefully designed to support the development of students' statistical reasoning.

OUR CONTEXT

Our introductory statistical literacy course is designed to prepare everyone, regardless of their statistical background, to develop the skills to be *critical consumers* of statistical information. Our definition of statistical literacy is based on Gal's framework (2002), with Gal's knowledge elements augmented to include heuristics and fallacies to highlight the need for a willingness to challenge prior beliefs (Budgett & Pfannkuch, 2010). We mainly focus on what Gal (2002) refers to as 'reading contexts', which arise when people encounter statistical and data-based information in their daily lives, including reading, watching, or listening to the news media.

Throughout the course, students evaluate statistical information across a wide range of sources, including TV clips, radio excerpts, social media posts, news media articles, technical reports, and academic articles. The focus is on conceptual understanding, which is facilitated through a broad range of in-class activities during lectures and discussion-based group work in tutorials. The learning goals are designed so that students in the course can move towards adopting a critical stance in response to news and popular media articles reporting on statistically based studies (see Table 1).

Table 1. Statistical literacy course topics and learning goals.

Topic	Learning goals
Introduction to media reports	• Identify statistical messages in news media articles • Consider how measures may be defined • Apply guidelines and "worry questions" to news media articles reporting on statistically based studies
Surveys and polls	• Acknowledge sources of variability that arise in polls and surveys • Use a 'rule-of-thumb' to estimate the margin of error for single poll percentages and differences between poll percentages • Apply guidelines and "worry questions" to news media articles reporting on polls and surveys
Experimentation	• Distinguish between observational studies and experimental studies • Identify good practice for both types of study • Identify problems that can arise for both types of study • Apply guidelines and "worry questions" to news media articles reporting on observational studies and experimental studies
Risk	• Identify and distinguish between the various ways in which risk can be communicated • Calculate and interpret relative risks and account for uncertainty • Identify situations in which we may be misled about risk
Media reports	• Interpret and critique graphical representations of information • Acknowledge sources of variability that arise in measures such as means and risks

	• Identify situations in which we may be misled about statistics, including graphical representations, communicated in the media
Statistical reasoning	• Identify situations in our daily lives where statistical information is present • Be aware of specific heuristics and the errors in thinking that can arise from them

Content and assessment are designed around the following learning outcomes for the course :

- Critically evaluate a statistically-based study and communicate and justify judgment on the study
- Calculate and/or interpret estimates and confidence intervals and assess claims made in the media
- Describe important characteristics of well-conducted polls, surveys, experiments and observational studies and identify potential biases
- Identify and discuss probabilistic and statistical concepts and reasoning embedded in everyday contexts
- Work effectively and collaboratively with others

Together, these learning outcomes target important characteristics of statistically literate individuals and facilitate the development of attributes and capabilities we wish to foster, in order to empower today's graduates to participate meaningfully in society.

While most of our students are enrolled in a Bachelor of Science degree, or in a conjoint degree in which they pursue two undergraduate bachelor's degrees at the same time, one of which is a Bachelor of Science, we also cater to students outside of this disciplinary area (see Table 2).

Table 2. Percentage of students by discipline (including conjoint) – last 5 years.

Cohort *class size*	Science %	Arts %	Commerce %	Other %
2020 *(n=126)*	75	19	6	0
2021 *(n=153)*	79	14	3	4
2022 *(n=121)*	67	16	8	9
2023 *(n=133)*	65	15	5	15
2024 *(n=120)*	68	17	4	11

Between 20 and 30 per cent of the students enrolled in a Bachelor of Science are pursuing a major in statistics, with the remainder majoring in disciplines such as psychology, mathematics, computer science, and environmental science. As a result, the course serves a diverse cohort, with only a minority intending to pursue a career in statistics or data analysis. Because the course is designed to develop statistical literacy skills, all students, regardless of their major, should recognise its importance for informed and active participation in society.

ASSESSMENT

Assessment is a cornerstone of effective teaching and learning. The GAISE guidelines highlight that "assessments must focus on understanding key ideas, and not just on skills, procedures, and computer answers" (Carver et al., 2016, p. 21). Assessment tasks in our statistical literacy course have been described previously (Budgett & Pfannkuch, 2010), with the first assignment designed to address the course learning outcome: *Critically evaluate a statistically-based study and communicate and justify judgment on the study*. Students are presented with a technical report on a sample survey or opinion poll, and a related news media article. They are asked to evaluate the study, communicate and justify their judgment on it, and challenge any claims made in the news media article on the study. We aim to select articles that have meaning and relevance to "… elicit students' critical reading, reflective thinking, and communication of thoughtful opinions" (Gal, 2023, p. 323).

To evaluate the study, students use a set of "worry questions" to prompt their thinking (Gal, 2002), which are structured within Utts' (2015) seven critical components framework. Through a

process of answering questions such as "How was the sample selected?" and "What was the wording of the questions?", students engage in a systematic and more critical examination of how the opinion poll or sample survey was carried out. The "worry" questions serve as prompts that help students to move beyond a surface-level interpretation and to develop reasoned judgments about the study. The most challenging part for students is harnessing their own personal knowledge or identifying issues that perhaps were not addressed by the report authors. However, some generate insightful comments on issues such as potential biases and possible hidden agendas. After completing their analysis, students are assessed on their ability to provide a reasoned opinion on the study. Some students struggle to synthesise and justify a coherent judgment, highlighting that both analysis and synthesis of statistical information are two interconnected components of statistical literacy, both of which need to be assessed. Students then consider the news media article reporting on the study and determine whether the claims are justified.

Students provide written responses to the three elements outlined above and submit them to the learning management system Canvas for marking. In the 2023 iteration of the course, we were alerted to several submissions that appeared to have been generated using AI. These were easy to spot, with some submissions containing overly generalised phrasing lacking contextualisation and others featuring fabricated sources or unusual language. While undoubtedly creating opportunities within the realm of higher education (Hoernig et al., 2024), there are justifiable concerns that generative AI (Gen-AI) tools can pose a threat to academic integrity. Traditional forms of assessment, whether on paper or digital, may no longer be fit for purpose if one cannot determine the legitimacy of an individual's submission. Particularly relevant for students in a statistical literacy course is "their ability to discuss or communicate their reactions to […] statistical information, such as their understanding of the meaning of the information, their opinions about the implications of this information, or their concerns regarding the acceptability of the given conclusions." (Gal, 2002, p. 3). While we recognise the potential benefit of AI for brainstorming and idea generation, it is important to reflect critically on the output of Gen-AI tools and to edit, revise, and restructure to convey personal and unique understandings and insights on a given topic. We are not benefiting our students by denying them an authentic opportunity to express, discuss, and defend their personal reactions to statistical information presented in the media. Conscious that our assessment method needed to evolve to maintain relevance and credibility, we explored alternative approaches.

Oral Assessment

Oral assessments can enhance authenticity, provide deeper insights into student understanding, foster the development of transversal skills, and promote academic integrity (Ward et al., 2024). Although rare, there are precedents for oral assessment in statistics courses. Recognising the potential to provide a richer picture of student knowledge, especially when the focus is on conceptual understanding (e.g., Carver et al., 2016), Theobold (2021) implemented oral exams in an applied regression course, and Yanchenko (2025) administered an oral exam with non-native English speakers in an introductory statistics course. Both researchers shared insightful reflections on their experiences and welcomed the opportunity to probe students' understanding in ways not typically afforded by traditional written assessments. While Theobold's (2021) article was published in the pre-Gen-AI era, another of Yanchenko's (2025) goals was to motivate students to engage with the course content and to avoid the temptation to "circumvent the learning process by simply passing off their assignments to LLMs" (p. 2). Different forms of oral assessment presentation formats include poster presentations, viva voce exams, pre-recorded presentations, group presentations, and interactive oral presentations (Ward et al., 2024).

An IOA is defined as an efficient and effective form of authentic assessment that promotes skill development and employability, enhances overall student engagement and a personalised approach to learning and teaching, and preserves academic integrity (Sotiriadou et al., 2020). The IOA facilitates a "genuine, unscripted, synchronous conversation between an assessor and a student (or a group of assessors/students) around an authentic workplace scenario […] which can be performed face-to-face or online" (Shaeri et al., 2021, p. 105). IOAs are effective across disciplines such as nursing (Gulikers et al., 2004) and law (Martens et al., 2007), fostering genuine conversations in which students supply course concepts in industry and real-world scenarios. In IOAs, students can demonstrate their ability to

extend and synthesise disciplinary knowledge and skills in diverse contexts, showcasing their proficiency while also developing their conversational and communication skills. Academic misconduct can easily be identified in an IOA because, in a genuine and unscripted conversation, students cannot successfully extend and synthesise work that they have not undertaken themselves.

IOA in our context

Sharing motivations similar to those of both Theobold (2021) and Yanchenko (2025), we introduced an IOA element to the first assignment. We opted for an IOA because of our focus on 'reading contexts', and the idea that the IOA can be considered less formal than some of the other oral assessment presentation formats (Ward et al., 2024). We also anticipated the IOA to be particularly relevant to assessing statistical literacy skills. Through participating in a two-way conversation and defending their statistical thinking in response to instructor probing, students can demonstrate their statistical reasoning and critical understanding of texts regularly encountered in their daily lives. Although the original intent of the assessment remained unchanged – requiring students to critically evaluate a statistically-based study using the "worry questions", formulate a reasoned judgment, and challenge any misleading claims in an associated news media article – the mode of submission differed. Instead of submitting a written response, students were required to engage in a five-minute IOA with one of the two course instructors.

Sotiriadou and colleagues (2020) proposed an evidence-based model supporting six characteristics they believe are necessary in designing an IOA to successfully achieve the objectives of student engagement, employability, and academic integrity. We adapted this framework to fit our own context (Table 3).

Table 3. IOA characteristics (Sotiriadou et al., 2020, p. 2144) and local context illustrations.

Characteristic	IOA 2024 implementation strategies
Scaffolding and support Demonstrates the scope of the assessment and how the assessment works	IOA briefings – two 15-minute sessions in lectures • Opportunities to ask questions Tutorial session – one 50-minute session • Modelling (IOA video example) • Collective marking of video IOA using the rubric • IOA practice with a tutorial group
Scenario-based context Enables students to see relevance to future work-related or social situations	Given the diverse backgrounds of our students (Table 2), we focused on highlighting the relevance of the IOA within a possible future social setting.
Aligned to programme Articulates alignment with programme learning outcome/s	Given the diverse backgrounds of our students (Table 2), we focused on highlighting the alignment of the IOA to the University of Auckland graduate profile attributes, which apply across all programmes.
Learning outcomes Articulates alignment with course assessment learning outcomes	We focused on explicitly linking one or more of the course learning outcomes regarding both statistical literacy characteristics and capabilities and graduate attributes.
Accessible and equitable Endeavours to meet the requirements of accessible and equitable delivery	We did what we could to ensure that the IOA was accessible and equitable for all.
Professionally focused Offers assessment tasks reflective of disciplinary workplace roles	Given the diverse backgrounds of our students (Table 2), we focused on the importance of statistically literate citizenship within a possible future social setting.

IOA implementation

To support students in their IOA preparation, we conducted in-class briefings that introduced the IOA and outlined the rationale for its implementation. We created a video example that students and instructors collectively marked in a tutorial session using the marking rubric. Students were encouraged to practice an IOA with tutorial group members to familiarise themselves with the process (Table 3).

We allocated two days at the end of the first half of the semester to conduct the IOAs. Students scheduled their appointment via the learning management system Canvas and had the option to attend face-to-face or on Zoom. Of the 120 students enrolled in the course, 115 completed an IOA, with the two course instructors – the authors – dividing the sessions between them. Each IOA was marked upon completion using a rubric shared with the students in advance. All sessions were recorded, and following their IOA, students were provided with an audio file of their assessment, their rubric-aligned marks, and personalised feedback. Students were allowed to contest their marks. To do so, they were required to review their recording, consult the rubric and feedback, and provide a rationale for their appeal. Students were also asked to complete a questionnaire to gauge their perceptions and experiences of the IOA.

Feedback from the students was overwhelmingly positive. Of the 72 students (63%) who completed the questionnaire, 81% reported feeling well-prepared for their assessment, and 89% expressed satisfaction with the information and support they received. A majority (87%) agreed that the IOA reflected real-world relevance and authenticity, although a smaller proportion (63%) perceived the scenario as being one they might realistically encounter. Encouragingly, 87% agreed that the IOA contributed to the development of skills that would be valuable beyond their university studies.

REFLECTIONS

As with any changes to established assessment approaches, careful preparation and planning were essential. We were fortunate to benefit from the expertise of learning designers from the Learning and Teaching Design Team at the University of Auckland, who facilitated a connection with a leading researcher in interactive oral assessment. This collaboration proved invaluable, offering significant guidance and support throughout the design, development and implementation of our first IOA.

Introducing the IOA in our introductory statistical literacy course for the first time required considerable effort, particularly in planning, coordination and resource development. Much of our preparation was ‘just in time’ as we negotiated the practicalities and complexities of designing and delivering a new assessment format whilst simultaneously ensuring adequate support for our students. We were transparent with our students, openly acknowledging that this was our first time using the IOA and that we had much to learn. We also conveyed our enthusiasm and expressed our hope that the students would find the experience enjoyable and engaging. Despite the challenges, the process proved to be enjoyable and rewarding for both instructors and students.

CONCLUSION

Introducing the IOA in our introductory statistical literacy course was both a challenging and rewarding experience requiring adaptability, collaboration, and commitment to student support. We believe that transparency with students, together with a shared enthusiasm for the process, helped to foster an engaging and collegial environment. The shift from traditional written assessments to an oral approach aligned with our goal of enhancing students’ communication skills, particularly in contexts involving the evaluation and judgment of texts such as news media articles. A key feature of the original assessment was the use of “worry” questions, which encouraged students to engage deeply and critically with a statistical report. The IOA provided a valuable platform for students to articulate their understanding and express their opinions and judgment on the study and the way in which it was presented in the news media. We believe that the IOA supports the development of essential skills for navigating data in everyday life. By encouraging students to interpret and evaluate statistical information and communicate their reactions, the IOA can empower students to express their ideas with clarity and confidence. These abilities – critical evaluation and effective communication – are two interrelated and foundational components of statistical literacy (Gal, 2002).

ACKNOWLEDGMENTS

The authors sincerely thank Heera Kim and Anne McKay, learning designers from the Learning and Teaching Design Team at the University of Auckland, New Zealand and Danielle Logan-Fleming of Torrens University, Australia. Their expertise and guidance were instrumental in the development and implementation of this work.

REFERENCES

Budgett, S., & Pfannkuch, M. (2010). Assessing students' statistical literacy. In P. Bidgood, N. Hunt & F. Jolliffe (Eds.), *Assessment methods in statistical education: An international perspective* (pp. 103–121). Wiley. https://doi.org/10.1002/9780470710470.ch9

Carver, R., Everson, M., Gabrosek, J., Horton, N., Lock, R., Mocko, M., Rossman, A., Roswell, G. H., Velleman, P., Witmer, J., & Wood, B. (2016). *Guidelines for Assessment and Instruction in Statistics Education (GAISE) College Report.* American Statistical Association. https://www.amstat.org/docs/default-source/amstat-documents/gaisecollege_full.pdf

Engel, J., & Ridgway, J. (2023). Back to the future: Rethinking the purpose and nature of statistics education. In J. Ridgway (Ed.), *Statistics for empowerment and social engagement: Teaching Civic Statistics to develop informed citizens* (pp. 17–36). Springer. https://doi.org/10.1007/978-3-031-20748-8_2

Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibility. *International Statistical Review, 70*(1), 1–51. https://doi.org/10.2307/1403713

Gal, I. (2023). Critical understanding of Civic Statistics: Engaging with important contexts, texts, and opinion questions. In J. Ridgway (Ed.), *Statistics for empowerment and social engagement: Teaching Civic Statistics to develop informed citizens* (pp. 323–343). Springer. https://doi.org/10.1007/978-3-031-20748-8_13

Gulikers, J., Bastiaens, T., & Kirschner, P. (2004). A five-dimensional framework for authentic assessment. *Educational Technology Research and Development, 52*(3), 67–86. https://doi.org/10.1007/BF02504676

Harari, Y., Harris, T., & Raskin, A. (2023, March 24). Opinion: You can have the blue pill or the red pill, and we're out of blue pills. *The New York Times*. https://www.nytimes.com/2023/03/24/opinion/yuval-harari-ai-chatgpt.html

Harford, T. (2021). *The data detective: Ten easy rules to make sense of statistics.* Riverhead Books.

Hoernig, S., Ilharco, A., Pereira, P. T., & Pereira, R. (2024). *Generative AI and higher education: Challenges and opportunities.* Institute of Public Policy Lisbon. https://www.ipp-jcs.org/wp-content/uploads/2024/09/Report-AI-in-Higher-Education-IPP-1.pdf

Klein, E. (2023, March 12). Opinion: This changes everything. *The New York Times*. https://www.nytimes.com/2023/03/12/opinion/chatbots-artificial-intelligence-future-weirdness.html

Lanier, J. (2023, April 20). Annals of Artificial Intelligence: There is no AI. *The New Yorker*. https://www.newyorker.com/science/annals-of-artificial-intelligence/there-is-no-ai

Martens, R., Bastiaens, T., & Kirschner, P. A. (2007). New learning design in distance education: The impact on student perception and motivation. *Distance Education*, *28*(1), 81–93. https://doi.org/10.1080/01587910701305327

Piper, K. (2022, November 29). The Highlight: AI experts are increasingly afraid of what they're creating. *Vox*. https://www.vox.com/the-highlight/23447596/artificial-intelligence-agi-openai-gpt3-existential-risk-human-extinction

Shaeri, S., Logan, D., & Krautloher, A. (2021). Evaluating competency development using interactive oral assessments. In *Research in Engineering Education Symposium and Australasian Association for Engineering Education Conference* (pp. 103–112). Engineers Australia. https://search.informit.org/doi/10.3316/informit.345915443844801

Sotiriadou, P., Logan, D., Daly, A., & Guest, R. (2020). The role of authentic assessment to preserve academic integrity and promote skill development and employability. *Studies in Higher Education, 45*(11), 2132–2148. https://doi.org/10.1080/03075079.2019.1582015

Theobold, A. S. (2021). Oral exams: A more meaningful assessment of students' understanding. *Journal of Statistics and Data Science Education, 29*(2), 156–159. https://doi.org/10.1080/26939169.2021.1914527

Utts, J. (2015). *Seeing through statistics* (4th edition). Cengage Learning.

Wallace-Wells, D. (2023, March 27). Opinion: AI is being built by people who think it might destroy us. *The New York Times*. https://www.nytimes.com/2023/03/27/opinion/ai-chatgpt-chatbots.html

Ward, M., O'Riordan, F., Logan-Fleming, D., Cooke, D., Concannon-Gibney, T., Efthymiou, M., & Watkins, N. (2024). Interactive oral assessment case studies: An innovative, academically rigorous, authentic assessment approach. *Innovations in Education and Teaching International, 61*(5), 930–947. https://doi.org/10.1080/14703297.2023.2251967

Yanchenko, E. (2025). Oral exams in introductory statistics class with non-native English speakers. *Teaching Statistics*. https://doi.org/10.1111/test.70015