**REshaping Assessment Excellence: COVID-19 and the New Now at the University of Bath**

**Davenport, J.H.**1 and Crick, T.2

1 - University of Bath, UK; 2 – Swansea University, UK.

**Abstract**

While there has been a great deal of debate in academic circles about the future of assessment, and the ideal form of it, actual change has been far more limited. “While faculty and students alike will not stray too far from a computer as they go about their daily business, it is still the norm for examinations to be conducted using pen and paper”. Will the COVID-19 pandemic be seen by history as the event that changed this, and actually allowed — or indeed, empowered — universities to adopt a broader range of assessment techniques in practice? We frame this using the University of Bath as an academic case study through the pandemic, reflecting on the recent period and the “new now”.

**Introduction**

The impact of the COVID-19 pandemic on the UK higher education sector has been significant (and indeed, ongoing), for both academic faculty (Watermeyer et al., 2021; Watermeyer et al., 2021) and professional services staff (Watermeyer et al., 2021; Watermeyer et al., 2022), inhibiting or curtailing all university activities. Whilst the pandemic has precipitated significant change in learning, teaching and assessment over the recent period (Crick, 2021), it will be interesting to see how sustained and far-reaching this becomes. It is hard to separate learning and teaching from assessment, as assessment drives much student motivation; some wider questions about student motivation and resilience during the pandemic are explored in Crick et al. (2022). Though there has been much debate about for many years the various forms of assessment, and the role of technology in assessment, actual change has been slow pre-COVID, as evidenced in this statement (Williams and Wong, 2009).

Indeed, the final examination is a university institution that would appear to be off-limits as far as innovation is concerned. To put this into context, while faculty and students alike will not stray too far from a computer as they go about their daily business, it is still the norm for examinations to be conducted using pen and paper. Does this imply, therefore, that some element of modern learning theory might be sacrificed if it were abandoned in favour of some alternative instrument? Or, given it is still the most commonly administered summative assessment instrument in universities today, is there some other special intrinsic value attached to a closed-book, invigilated exam that justifies its continued use?

The authors wonder whether the COVID-19 pandemic has not been the trigger that puts such innovation within the scope of debate.

**Models**

**Trad-C** A university-managed invigilated[[1]](#footnote-1) examination in an “examination hall”, with no technology or other resources except the question paper (and possibly tables etc.).

**Trad-OR** As above but students can bring in specified paper resources: often called “open book”. That specification could be very precise: “A clean copy of the fifth edition of …”, or as vague as “a binder of own notes”. In practice invigilators find this hard to enforce in a large examination., e.g. that a textbook has no annotations. JHD had used this on XX10190 (Betteridge *et al.*,2019), essentially to deter students from memorising the details in the book. But many years ago he offered MSc students the choice, and they voted for Trad-C (probably out of familiarity).

**Trad-OU** As above with no restrictions on what can be brought it (on paper).

**TakeHome-OU** In a take-home examination, the student is given the question paper, and has to bring the answer back later (generally 24-hours). Bengtsson (2019) is a useful survey of these, and has two key points.

1. It is concluded that take-home exams may be the preferred choice of assessment method on the higher taxonomy levels because they promote higher-order thinking skills and allow time for reflection.
2. Due to the obvious risk of unethical student behaviour, take-home exams are not recommended on the lowest taxonomy level.

JHD had experimented with this in the past for CM50209 Cybersecurity.

**Interim** Use a Virtual Learning Environment to deliver an examination paper, and collect answers. There are no technological constraints on the help students could acquire.

**Electronic-C** A university-managed examination, generally using a specific software platform (Bath used Inspera, but the precise choice is probably irrelevant) for delivery and submission. The students are not allowed to use any other resources, but there are wide variations on how this is enforced, from an honour system, through restricted browsers[[2]](#footnote-2) to full AI-based monitoring. Though it has been sold as a panacea, AI-based monitoring has its limitations (New York Times (Kashmir Hill), 2022) and may be illegal in some jurisdictions (JISC, 2022).

**Electronic-OU** As above, but the students are allowed to use any Internet resources. The wording here is not standardised, but the intention is that the students can consult internet resources, but not people. This raises unsolved questions around “intelligent”’ resources, notably those that can write University-level essays (Sharples, 2022), or even answers to programming exercises (Finnie-Ansley et al., 2022).

**University of Bath Timeline**

The University of Bath has operated a semester system since 1997 for practically all programmes, with examination assessments in, essentially, January and May. Before COVID-19, the only option for an examination was Trad-C or Trad-O, nearly all of two hours duration. Though this was not always possible, the aim was that a student should only have one examination per day. It was possible to disguise a **TakeHome-OU** examination as “coursework”.

**May 2020** Given the timing, Interim was the only practicable solution. Since many solutions were “at home” across the world, the examinations were still aimed at taking two hours, but all students were given a 24-hour window in which to do them. Very few staff had experience in “open book” examinations, and certainly not when the whole Internet was an open book. Now we needed to have one examination per day, which stretched the examination period.

**January 2021** We moved to Electronic-O. Because students were in different time zones, it was felt that a fixed start time was impossible, so there was still a 24-hour window. Mathematical Sciences took the option to insist that students only had three hours (conceived of as a 2 hour exam plus an hour for administration) to complete the examination from starting the process. But the students could still choose their start time (based on their home time zone) as long as the exam was done in the 24-hour window.

**May 2021** Based on the success of the Mathematical Sciences limited time experiment, and probably because staff now had more experience of setting open-book exams and getting the time requirements roughly right[[3]](#footnote-3), many more departments moved to the three-hour limit.

**January 2022** Now that students were “largely expected” to be at Bath, the University kept to three-hour (still thought of as 2+1) examinations, but now fixed the time, rather than allowing a 24-hour window.

**May 2022** Following very substantial pressure by the academics, the university allowed some Trad-C examinations in first-year subjects.

**Academic Misconduct**

Dickinson (2022) reports a small (N=900) survey (Alpha Academic Appeals, 2022)[[4]](#footnote-4) of UK students.

The numbers suggest that 1 in 6 students in the UK have cheated in online exams this academic year. Over half of those surveyed knew people who had cheated in online assessments. Almost 8 out 10 believed that it was easier to cheat in online exams than in exam halls, and the methods for cheating were often laughably rudimentary – including calling or messaging friends for help during the exam, using google to search for answers on a separate device, or asking parents to read through answers prior to submission.

The University of Bath has certainly experienced an increase in detected use of “unfair means” (the phrase used in Regulations). There is probably a larger increase in undetected use of unfair means. JHD has sat on judgement panels, and his subjective view would be that much of this has been extempore abuse, as students get tempted in the stress of the examination, rather than pre-planned (as cheating in **Trad-C** examinations has to be).

There is a particular concern over the use of transformer-based essay generators (Sharples, 2022). “Plagiarism software will not detect essays written by Transformers, because the text is generated, not copied. A Google search of the essay shows that each sentence is original”.

**Practitioner Experience**

In this period, JHD has been teaching two modules with examinations.

**CM30070—Computer Algebra** A module aimed at senior-year mathematicians and computer scientists, based on the author’s updated draft version (given to students) of Davenport et al. (1988). This had always had a **Trad-C** examination. Moving this to **Electronic-OU** wasn’t easy, and the first attempt was (rightly) criticised by the students as lengthy. One student notified JHD, during the examination, that one question had been posted on Chegg, and JHD immediately registered here, but was unable to see any offers or solution. The question was probably unanswerable by anyone who had not read the updated draft. JHD’s experience, and the move to a fixed three-hour (2+1) period meant that the examination in January 2022 went pretty well. In an ideal world, he would probably choose **Electronic-OR.**

**CM50209–Cybersecurity** A module aimed at MSC Computer science students, but also taken by others as well. This was mostly (80%) assessed by coursework, but also had a 20% **Trad-C** examination. This was replaced by **Electronic-OU**, very successfully. JHD would like it to stay that way: it is far more authentic as a Cybersecurity experience.

**Student Voice**

In the UK system, the “National Student Survey” (reference?) is a powerful instrument, for better or for worse. It is taken by final year undergraduate students in February-March of their final year, so after Semester 1 examinations in their final year, but before the last examinations and final results. In the case of Mathematical Sciences this will be a mixture of four cohorts (and the odd deferred student). For the 2022 survey, this will be as in the table.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Cohort | 2021/22 | 2020/21 | 2019/20 | 2018/19 | 2017/18 |
| Bsc | As above | As Above | Trad+Emerg | — | — |
| Bsc with Placement | As above | On placement | Trad+Emerg | Traditional | — |
| MMath | As above | As Above | Trad+Emerg | Traditional |  |
| MMath with Placement | As above | As Above | On placement | Traditional | Traditional |

There is no way of associating NSS comments with cohorts, which makes this less than ideal for our purposes. In 2022, one question asked was “What changes introduced this academic year, due to the Covid-19 pandemic, would you like to see maintained on your course for the benefit of future students?”. This had 69 responses in all, of which 22 were about examinations. These were the following (exam portions extracted):

1. Online exams which test knowledge as opposed to testing your memory. These test your understanding and application of your knowledge which is more useful for the real-world, as opposed to just remembering facts and figures;
2. The change of exams to more knowledge based than memory is refreshing;
3. Online exams - the exams work better when they use problems to check your understanding instead of your memory;
4. I think online exams via Inspera are a better way to test understanding on content rather than recall of definitions, etc.;
5. Open-book exams (7 times);
6. Open book exams. When done well, these test students at a much higher level by reducing the amount of memorisation that earns you marks. You must have a much increased understanding of the course to attain a first in an open book exam compared to the rote memorisation required for in-person exams;
7. Open book exams, previous are heavily based around memorising proofs, which do not test mathematical ability;
8. Open book exams are better for testing understanding, but in general are much harder to pass and should be altered to ensure a pass mark is still as attainable as it was for closed book;
9. Online exams, so exams aren't memory tests;
10. Online exams with varied start times;
11. Online exams which test understanding of concepts more than memory;
12. Online exams rather than in person;
13. I think open book exams are much better in terms of testing understanding rather than just memory;
14. I think open-book exams are a step in the right direction (away from memorisation, so that course/concept understanding can be tested more thoroughly, as I believe memorisation is not as important as it once was in most fields of study) but the removal of definition/'easy' or 'safe' questions can make the exam more difficult overall - I would like to see open-book exams but maintained at a closely similar level of difficulty/passability to the previous written exams;
15. Assessment style - but this needs a longer time frame than 2 hours;
16. As exams move back to in-person, have less questions based on theorems and proofs that can be copied from lecture notes as seen in online format of exams introduced due to COVID;
17. I absolutely would not want to keep online exams, in person is much better due to an abundance of cheating.

We can note that online examinations received overwhelmingly positive comments, but number 17 is a strong dissenter.

Computer Science students have a similar breakdown to the above in terms of demographics. The 13 comments on examinations (out of 45 total on Covid measures) were

1. Certain exams are more relevant in open-book, online format, but not all.
2. Online exams. I think it is a much more effective method of assessment! A lot of the anxiety is alleviated, so I feel I am able to perform much better. Furthermore, the fact that they are open book means that we are forced to show an in-depth understanding, rather than simply memorising facts.
3. Online exams which are based more on application of knowledge than memorization.
4. Online exams (x3);
5. Normalised methods of examination across all courses.
6. I haven't liked online exams or lectures;
7. Open book examinations which are designed to assess understanding rather than memorization are far more reflective of a student's success in learning the course. Therefore, I think that open books exams should be here to stay.
8. Online, open book exams. Closed book, more strictly limited time exams are not representative of real world ability, more similar to a memory test in certain modules.
9. I personally think online exams work well as the exams tend to be less memory based. This is good as the Internet exists and memory is of less importance than problem solving. The online exams therefore have better questions.
10. I much prefer the online exams that test the understanding of knowledge rather than how well I have managed to memorise something.
11. online exams - where application is more applicable than theory (open book exam format).

**Conclusions**

It is too early to draw conclusions, and the single institutional case study of the University of Bath is far too small a sample. Behaviours at Bath during the pandemic certainly seem to be in line with much of the rest of the UK HE sector, which is reinforced by wider outcomes of five major empirical studies conducted by the second author and colleagues over the past two years (Watermeyer et al., 2021a; Watermeyer et al., 2021b; Watermeyer et al., 2021c; Watermeyer et al., 2022). The authors’ discussions with colleagues from across the world indicate that behaviours in computer science world-wide seem similar, though there is much more reliance on AI-based invigilation (Siegel et al., 2021).

All colleagues agree, though, that “assessment format” is now firmly on the table. As in so many other areas of “digital transformation” (Crick, 2021), it may be that the COVID-19 pandemic has done to education what CEOs and CIOs have failed to do in the general IT-using sector. However, this may end up being piecemeal. At Bath, the Mathematical Sciences Department has reverted to the *status quo ante*. Students have been told this.

• In person: you will sit the exams at fixed times on fixed days in a venue at the University. The exams will be invigilated.

• Closed-book: You will not be allowed to have any revision materials with you or any access to the internet. Your exam papers will be tailored to this setting.

**Why us?**

• It is fairer: all students sit the exam in the same circumstances with no vagaries of internet quality or space to work.

• It is fairer: we have had several reports of collusion in exams (people doing the exam together around the kitchen table).

• It is the best way to examine mathematics: the learning outcomes of your courses require that you demonstrate both knowledge and understanding and a closed book exam is the fairest and most reliable way to test this.

The authors would largely agree with the first two points, but both they and the vast majority of the student quoted above would disagree with the first *as a universal statement.* Is “one size fits all” back to haunt us?

**References**

Alpha Academic Appeals (2022). Press release on prevalence of cheating in online assessment, July 2022. <http://www.academicappeals.co.uk/news/05072022201747-press-release-on-prevalence-of-cheating-inonline-assessment--july-2022/>, 2022.

L. Bengtsson (2019). Take-Home Exams in Higher Education: A Systematic Review. *Educ. Sci*., 9:267–282, 2019.

Betteridge *et al.* (2019). Jack Betteridge, James H. Davenport, Melina Freitag, Willem Heijltjes, Stef Kynaston, Gregory Sankaran, and Gunnar Traustason. Teaching of computing to mathematics students: Programming and discrete mathematics. In Steven Bradley and Alexandra Cristea, editors, CEP ’19: *Proceedings of the 3rd Conference on Computing Education Practice*, pages 12:1–12:4, New York, NY, USA, 2019. ACM.

T. Crick (2021). “COVID-19 and Digital Education: A Catalyst for Change?” *ITNOW*, vol. 63, no. 1, pp. 16–17, 2021, <https://doi.org/10.1093/itnow/bwab005>

T. Crick, C. Knight, R. Watermeyer, & J. Goodall (2020). “The Impact of COVID-19 and “Emergency Remote Teaching” on the UK Computer Science Education Community,” in Proceedings of UK and Ireland Computing Education Research Conference (UKICER’20). ACM, 2020, <https://doi.org/10.1145/3416465.3416472>

T. Crick, T. Prickett, & J. Bradnum (2022). Exploring Learner Resilience and Performance of First-Year Computer Science Undergraduate Students during the COVID-19 Pandemic. Proceedings of the 27th ACM Conference on Innovation and Technology in Computer Science Education (ITiCSE’22), pages 519–525, 2022. https://doi.org/10.1145/3502718.3524764

J. Davenport, Y. Siret, & E. Tournier (1988). *Computer Algebra*. Academic Press, 1988.

J. Dickinson (2022). Did 1 in 6 students cheat in online assessments this year? https://wonkhe.com/blogs/did-1-in-6-students-cheat-inonline-assessments-this-year/, 2022.

Finnie-Ansley.J., P. Denny, B. Becker, A. Luxton-Reilly, & J. Prather (2022). The Robots Are Coming: Exploring the Implications of OpenAI Codex on Introductory Programming. ACE ’22: Australasian Computing Education Conference, pages 10–19, 2022.

New York Times [Kashmir Hill] (2022). Accused of Cheating by an Algorithm, and a Professor She Had Never Met. <https://www.nytimes.com/2022/05/27/technology/college-students-cheating-software-honorlock.html>, 2022.

Jisc/Uniwise. (2022). Online proctoring: panacea or problem? https://www.jisc.ac.uk/membership/stories/online-proctoring-panacea-or-problem, 2022.

M. Sharples (2022), New AI tools that can write student essays require educators to rethink teaching and assessment. <https://blogs.lse.ac.uk/impactofsocialsciences/2022/05/17/new-ai-tools-that-can-write-student-essays-require-educators-to-rethink-teaching-and-assessment/>.

A. Siegel, M. Zarb, B. Alshaigy, J. Blanchard, T. Crick, R. Glassey, J. R. Holt, C. Latulipe, C. Riedesel, M. Senapathi, Simon, and D. Williams (2021), “Teaching through a Global Pandemic: Educational Landscapes Before, During and After COVID-19,” in Proceedings of the 2021 Working Group Reports on Innovation and Technology in Computer Science Education (ITiCSE- WGR’21), 2021, <https://doi.org/10.1145/3502870.3506565>

R. Watermeyer, T. Crick, & C. Knight (2021a), “Digital disruption in the time of COVID-19: Learning technologists’ accounts of institutional barriers to online learning, teaching and assessment in UK universities,” International Journal for Academic Development, 2021, special issue on “Academic Development in Times of Crisis” <https://doi.org/10.1080/1360144X.2021.1990064>

R. Watermeyer, T. Crick, C. Knight, & J. Goodall (2021b), “COVID-19 and digital disruption in UK universities: afflictions and affordances of emergency online migration,” Higher Education, vol. 81, pp. 623–641, 2021, <https://doi.org/10.1007/s10734-020-00561-y>

R. Watermeyer, K. Shankar, T. Crick, C. Knight, F. McGaughey, J. Hardman, V. Suri, R. Chung, & D. Phelan (2021c), “’Pandemia’: A reckoning of UK universities’ corporate response to COVID-19 and its academic fallout,” British Journal of Sociology of Education, vol. 42, no. 5-6, pp. 651–666, 2021, <https://doi.org/10.1080/01425692.2021.1937058>

R. Watermeyer, C. Knight, T. Crick, & M. Borras Batalla. “‘Living at work’: COVID-19, remote-working and the spatio-relational reorganisation of professional services in UK universities,” Higher Education, 2022, <https://doi.org/10.1007/s10734-022-00892-y>

B. Williams and A. Wong (2009). The efficacy of final examination: A comparative study of closed-book, invigilated exams and open-book, open-web exams. *Br. J. Educ. Technol*., 40:227–236, 2009.

1. We use the British English word “invigilated” — the corresponding American word is “proctored”. [↑](#footnote-ref-1)
2. Note that restricted browsers in an invigilated computer room seems to be a pretty satisfactory

   solution from the academic integrity point of view, though it does have resource implications, and would clearly not have been viable with strict isolation measures. [↑](#footnote-ref-2)
3. It is under-appreciated how difficult this is: among the most valuable feedback from colleagues or external examiners, especially to relatively new lecturers (to the subject/level combination, even if experienced elsewhere) can be about time requirements. [↑](#footnote-ref-3)
4. One item from this not reproduced is “Of those students who admitted to cheating, only a very small minority — 5% — were caught by their institutions”. [↑](#footnote-ref-4)