

COMPSCI 289: Individual Seminar

HO YIN ALVIN CHEUNG (ACHE706)

A solid blue horizontal bar spanning the width of the slide, located at the bottom.

Introduction

Topic of the paper:

Cheat-resistant multiple-choice examinations using personalization [1]

Discipline of the paper: Computing Education

Reasons for developing the topic:

MC exams are susceptible to cheating

- High achievers are more likely to cheat [2] [3]
- Many existing cheating strategies in MC exams [1]

Consider the following DTD declaration of an element K and three instances of this K element:

`<!ELEMENT K (P?Q*R?S?T)>`

X. `<K><P/><P/><P/><Q/><R/><R/><S/><S/></K>`

Y. `<K><P/><Q/><Q/><Q/><S/><T/></K>`

Z. `<K><R/><S/><T/></K>`

Which of the element instances conform to the DTD declaration?

- A. X and Y only
- B. Y and Z only
- C. X and Z only
- D. All of X, Y, and Z
- E. None, or only or one of X, Y, and Z

Y, Z

Fig. 1: Collusion in MC exams

Source: Adapted from [1]

Introduction

Distinctions of this paper from associated topics [1]:

1. Addresses on assessments, not teaching
2. Integrating personalisation with exams
 - Few similar attempts in the past
3. Also addresses the pedagogical and administrative challenges



Fig. 2: a public examination
[used as decoration]
Source: Adapted from [4]

Introducing Personalised Tests

Three most common test personalisation methods [1]:

1. Randomising question parameters
2. Drawing questions from a pre-set question bank
3. Generating questions using macro scripts

Objectives

Objectives (research questions) of the paper [1]:

1. Attempting to create a framework for personalised MC exams
 - How feasible is the framework?
2. Finding out the technical, pedagogical, and administrative challenges specific to exam personalisation
 - Possible methods to overcome the challenges?



Fig. 3: someone taking a mathematics examination
[used as decoration]
Source: Adapted from [14]

Previous efforts of the Topic

Some other resources regarding similar topics:

1. Personalised computer-based tests [5]
2. Personalised online education [6]
3. The researcher's past research paper on plagiarism mitigation with personalisation [7]
4. QAA's guidance on addressing contract cheating [8]
5. Many recent works on statically analysing possible collusions [9], [10], [11], [12]

Challenges to Exam Personalisation

Administrative challenges [1]:

➤ The research took place at UoA

1. Cannot use digital exam due to large class size
 - Have to print and distribute physically
2. Cannot print version number on Scantron sheets
 - Have to encode the number as question choices
3. Need a pre-processing facility to convert the encoded version number

Challenges to Exam Personalisation

Technical challenges [1]:

➤ Types of MC questions

1. How to support multiple types of questions?
 - Extended mention on XYZ questions
 - (elaborated in pedagogical challenges)
2. How to avoid duplicate options?
3. How to effortlessly register the correct option?
4. How can we correct marking mistakes post-exam?

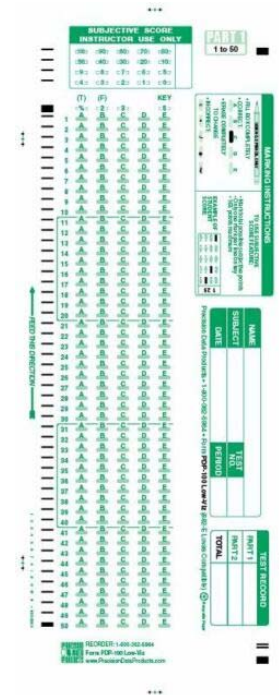


Fig. 4: A sample Scantron MC answer sheet
Source:
Adapted from [13]

Challenges to Exam Personalisation

Pedagogical challenges [1]:

1. How to ensure fairness of the exam?
2. How to create plausible distractors?
3. How can the generated exam cover most/all learning outcomes?
4. How to ensure the quality of the exam?

Framework for Exam Generation

Research framework: HTML with macros [1]

- Why use HTML?

Methods to accommodate teachers not wanting to program

- Allow creation of true/false questions

XML specification of the framework

- Samples of HTML templates and exam scripts

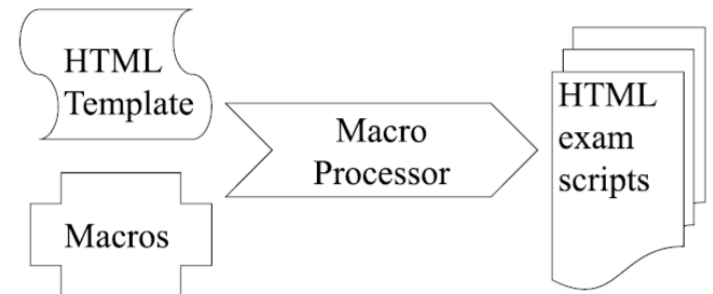


Fig. 5: An overview of the framework's HTML macro processor

Source: Adapted from [1]

Methodology of Research

Recalling the research questions [1]:

1. How feasible is to construct a generic framework (interface) to substantiate personalised examinations?
 2. What are the challenges introduced by exam personalisation?
- The responses are mostly positive

The system tried out in a graded in-class test

Methodology of Research

After the test: [1]

Researcher surveyed the students under trial anonymously on:

1. How resilient was the test against cheating?
2. Was the test fair?
3. Should other courses also use personalised tests?
4. How much did you like personalised tests overall?
 - The responses used a 1-5 point scale
 - Open-ended feedback also available



Fig. 6: a guy viewing some statistics
[used as decoration]
Source: Photo by Adeolu Eletu on Unsplash [15]

Methodology of Research

Staff surveys also offered [1]

Questions involved:

1. Comparing standard 4-version tests with personalised tests in terms of cheating
2. How much time did you spend to prepare personalised tests?

The research also compared the difference of grade distribution between using 4-version tests and personalised tests

A solid blue horizontal bar spanning the width of the slide, located at the bottom.

Initial Trial

First trial: a Compsci 3xx class

- Had slightly over 400 students
- Experiment conducted on a for-credit, in-class supervised exam

Exam script reviewed by ten staff members

- Two out of the ten staffs reviewed the source macro

Two other teaching assistants checked a generated sample of scripts

Initial Trial

Response rate of the post-exam survey: ~30% [1]

- Low response rate normal for undergraduate courses

- Summary of survey results

Found three questions having errors in answer generation post-exam

- Over 40% of students potentially marked wrong in at least one question
- Nine might be mismarked for all three erroneous questions

One mistakenly marked the wrong script ID

- Raised suspicion of collusion (found to be false)



Fig. 7: someone copying a neighbour's answers
[used as decoration]
Source: Adapted from [16]

Results from Staff Survey

Staffs attempted a standard 4-version test while trying to cheat [1]

- Found possible by identifying answer features
- Also possible to collude with big letters

Found personalised tests impossible to cheat unless allowing discussion

Staffs need extra time (2-3 times more) to develop questions for:

- Writing macros
- Designing pool of true/false question

Staffs unanimously agreed personalised tests reduce cheating

Further Trials

Rolled out in more CompSci papers [1]

- Similar class sizes as the first trial

The administration enthusiastically approved on personalised exams

Generally positive feedback from students over many courses

Comments focused on cheating reduction and positive learning

- Concerns about fairness among generated scripts

Found no evidence of statistically significant differences in grades

- Any difference possibly by student composition among various years

Issues from the previous trial persisted

- (Suspected) cheating cases found

Personal Comments

Pros:

1. The researcher conducted comprehensive trials on personalised tests
2. Completed some relevant side-investigations for the paper
3. Research questions mostly well-answered

Cons:

1. A few unclear expressions
 - Macro being a subset of parameter randomisation
 - “Same level of difficulty”?
2. Discrepancies of difficulty between versions remained unsolved

Possible Future Investigations

May need further trials in different faculties

- Existing runs only in Compsci

Should/can larger-scale exams (e.g. public exams) adopt personalisation?

How to reduce the difference of difficulties among the generated versions of the exam?



Fig 8: a dog typing in front of a computer [used as decoration]
Source: adopted from [17]

Conclusions

- The researcher-designed framework is generally successful
 - Instructors needed more time than usual to create customised exam
 - All interviewed staffs agreed that personalised exams reduce cheating
 - Students and the administration mostly approved the MC exam personalisation
- The research handled most of the challenges raised by personalisation
 - Used workarounds for exam version numbers
 - Instructors can correct marking errors after the exam
 - Addressed the challenge of creating plausible distractors by macros and XYZ questions
 - Still need future effort in tackling the unfairness caused by randomisation

References

- [1] S. Manoharan, "Cheat-resistant multiple-choice examinations using personalization," *Computers & Education*, vol. 130, pp. 139-151, 2019, doi: 10.1016/j.compedu.2018.11.007.
- [2] G. Yaniv, E. Siniver and Y. Tobol, "Do higher achievers cheat less? An experiment of self-revealing individual cheating," *Journal of Behavioral and Experimental Economics*, vol. 68, pp. 91-96, 2017, doi: 10.1016/j.socec.2017.04.005.
- [3] K. Ottaway, C. Murrant and K. Ritchie, "Cheating after the test: who does it and how often?," *Adv Physiol Educ*, vol. 41, pp. 368-374, 2017, doi: 10.1152/advan.00103.2016.
- [4] HKEAA. "The 19th issue of the HKDSE newsletter." HKEAA.
https://www.hkeaa.edu.hk/hkdse_newsletter/vol19/images/001.jpg (accessed Aug. 21, 2020).
- [5] C. Smaill, "The Implementation and Evaluation of OASIS: A Web-Based Learning and Assessment Tool for Large Classes," *IEEE Transactions on Education*, vol. 48, no. 4, pp. 658-663, 2005, doi: 10.1109/TE.2005.852590

References [cont.]

- [6] D. Weld, E. Adar, L. Chilton, R. Hoffmann, E. Horvitz, M. Koch, J. Landay, C. Lin and Mausam. (2012). Personalized Online Education — A Crowdsourcing Challenge. Presented at Association Advancement of Artificial Intelligence. [Online]. Available: <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.1052.1539&rep=rep1&type=pdf>
- [7] S. Manoharan, "Personalized assessment as a means to mitigate plagiarism," *IEEE Transactions on Education*, vol. 60, no. 2, pp. 112–119, 2017, doi: 10.1109/TE.2016.2604210.
- [8] The Quality Assurance Agency for Higher Education, "Introduction" in *Contracting to Cheat in Higher Education: How to Address Essay Mills and Contract Cheating*, 2nd ed, UK: QAA, ch. 1, sec. 2, pp. 3.
- [9] O. Wesolowsky, "Detecting excessive similarity in answers on multiple choice exams," *Journal of Applied Statistics*, vol. 60, no. 7, pp. 909–921, 2000, doi: 10.1080/02664760050120588.
- [10] A. Ercole, D. Whittlestone, G. Melvin and J. Rashbass, "Collusion detection in multiple choice examinations," *Medical Education*, vol. 36, no. 2, pp. 166–172, 2000, doi: 10.1046/j.1365-2923.2002.01068.x.

References [cont.]

[11] P. Richmond and M. Roehner, "The detection of cheating in multiple choice examinations," *Physica A: Statistical Mechanics and Its Applications*, vol. 436(Supplement C), pp. 418–429, 2016, doi: 10.1016/j.physa.2015.05.040.

[12] A. D'Souza and V. Siegfeldt, "A conceptual framework for detecting cheating in online and take-home exams," *Decision Sciences Journal of Innovative Education*, vol. 15, no. 4, pp. 370–391, 2017, doi: 10.1111/dsji.12140.

[13] buybazinga. "Scantron 882 E Lovas COMPATIBLE Testing Forms (50 Sheet Pack)". Ebay. <https://www.ebay.com/c/1242034023> (accessed Aug. 22 2020).

[14] Stashing Dollars. "How to Save Time and Money with CLEP Exams". Stashing Dollars. <https://stashingdollars.com/clep-study-guides/> (accessed Aug. 22 2020)

[15] Adeolu Eletu. [No title]. Unsplash. <https://www.citethisforme.com/guides/ieee-with-url/how-to-cite-a-online-image-or-video> (accessed Aug. 22 2020)

References [cont.]

[16] S. Campbell. "Foreign students FOUR TIMES more likely to cheat as officials catch 50,000 exam fraudsters". Express. <https://www.express.co.uk/news/uk/630829/essay-writing-service-foreign-students-cheating-exams-university/> (accessed Aug. 23 2020)

[17] LIHKG. [No title]. LIHKG. <https://cdn.lihkg.com/assets/faces/dog/itdog1.gif/> (accessed Aug. 23 2020)