9th International Conference on Higher Education Advances (HEAd'23)
Universitat Politècnica de València, València, 2023
DOI: http://dx.doi.org/10.4995/HEAd23.2023.16371

When a test-taking strategy is better? An approach from the paradigm of scheduling under explorable uncertainty

Cristóbal Alfredo Mauricio, Sebastián Davila-Gálvez, Óscar Carlos Vásquez
University of Santiago of Chile (USACH), Faculty of Engineering, Program for the Development of Sustainable Production Systems (PDSPS), Chile, University of Santiago of Chile (USACH), Faculty of Engineering, Industrial Engineering Department, Chile.

Abstract

In this article, we adopt the paradigm of scheduling under explorable uncertainty to explore test-taking strategies to solve standardized tests in terms of maximizing the correct questions answered. From this approach, a test taker considers a number of questions and has the possibility to read in order to obtain the difficulty of the question. Later, he/she has the option, for example, to answer the question or to skip the one that seemed difficult and read the next question in the test. Specifically, we state the problem definition by considering two test-taking strategies, formulate and implement a mathematical model, and generate computational experiments in order to determine the dominance of one strategy over another. The results show that the dominance depends directly on the design of the test and the maximum time to perform it, so knowing these parameters allow us to provide algorithmic insights to address this problem.

Keywords: *Test-taking strategies;scheduling; standarized test; optimization.*

1. Introduction

In the education setting, there are different types of standardized tests. These consider a wide range of knowledge from the foreign language proficiency tests (e.g., Test of English as a Foreign Language (TOEFL), Test of English for International Communication (TOEIC), among others) to the mandatory admission test such as the Graduate Management Admission Test (GMAT) in the United States of America or the Prueba de Acceso a la Educación Superior (PAES) in Chile (DEMRE, 2023), whose results define the position in the admission ranking of the university degree to be studied. In general, the students prepare to take these standardized tests over many months, attending some specialized courses where several test-taking strategy recommendations are provided to achieve the best possible result. In particular, a test-taking strategy from a scheduling point of view distinguishes a series of decisions to be made by the test-taker in sequence when faced with the questions in an assessment. For instance, the test-taker would read and answer all the questions in order of appearance or only answer the questions if he/she is sure of the correct answer. Considering that assessments typically have a limited time frame for completion, the choice of test-taking strategy can significantly impact the test results. Rapidly skipping challenging questions may allow test-takers to attempt more questions and answer the easier ones first. Still, it may also require them to revisit the skipped questions later, thus increasing the overall test completion time. Therefore, there exists a threshold where it may be more advantageous to respond immediately, regardless of the difficulty of the questions, rather than postponing difficult questions and answering easier ones first.

In literature, several studies have been carried out in different areas. In psychology, most of the strategies have been studied before taking the test in order to reduce the anxiety of the students (Theobald, Breitwieser, & Brod, 2022), and it has been found that the application of these, in conjunction with the application of other strategies has a direct correlation with student achievement in college (Viet, 2022). From a practical perspective, classic books such as "Test-Taking Strategies" (Kesselman-Turkel & Peterson, 1981) expose in a practical way several strategies for different types of tests, such as strategies for multiple-choice tests, true-false tests, matching tests, vocabulary tests, number tests, etc. However, the choice of these strategies is vaguely explained. Finally, from an algorithmic point of view, the decision of which question to answer first can be studied as a problem of allocating limited resources to tasks over time such as (Wang & Zhang, 2006), which presents an iterative mathematical model applied to a general case as well as a theoretical framework using nonlinear optimization, considering a probabilistic perspective previously known by the test-taker.

In this paper, we adopt the paradigm of scheduling under explorable uncertainty (Retsef, Magnanti , & Shaposhnik, 2018) (Dürr, Erlebach, Megow, & Meißner, 2020) (Dufossé, Dürr, Nadal, Trystram, & Vásquez, 2022) to explore test strategies to solve standardized tests and to analyze under what circumstances the application of these strategies will generate an optimal result. From this approach, a test-taker considers a number of questions (set of parameters) and has the possibility to read (to make a query) in order to obtain the difficulty of the question. Later, he/she has the option, for instance, to answer the question or to skip the one that seemed difficult and read the next question in the test. Clearly, a compromise has to be found between the number of questions to be skipped and the test result obtained. Formally, we state the problem definition by considering two test strategies, formulate and implement a mathematical optimization model and generate computational experiments in order to determine dominance properties of the strategies to obtain the best test result.

2. Problem definition and formulations

2.1. Problem description

Consider a test with N questions, of which $M \leq N$ are easy questions, and N-M are hard questions. P is defined as the set of questions $\{1, \dots, N\}$ where each question $i \in P$ has a reading time r_i and an answer time p_i. Each question appears in the test follows a particular sequence S, e.g., $S := \{\mathrm{h}, \mathrm{h}, \mathrm{e}, \mathrm{h}, \mathrm{e}, \mathrm{e}, \dots, \mathrm{h}\}$, where h denote the hard questions and e the easy questions. These questions are differentiated by their resolution time, where the time to answer a hard question p_h will be longer than the time to answer an easy question p_e. Additionally, there is a deadline $\overline{d}$ to finish the test. We assume that the test-taker will know the number of total questions (N) and the easy questions (M) in the test.

We consider two test strategies for answering the test, which are described below:

- Read and answer (RA) strategy: This test-taking strategy consists of reading and answering the question immediately, regardless of the difficulty of the question.
- Read, skip, and answer (RSA) strategy: This test-taking strategy consists of reading in an exploratory way and analyzing the difficulty of the question. If the question is easy, it is answered immediately, but if it is hard, then it is skipped and answered once all the easy questions have been answered. Skip one question $i \in \{1, \dots, N\}$ has an associated time cost ℓ_i which will consist of re-reading the hard question before answering it.

This problem aims to study the best strategy to answer as many questions as possible into the deadline $\overline{d}$. The choice of a strategy when facing a test can drastically affect the result. This result depends on several factors; the test design is the most important to consider (such as sequence, read, response, and re-read times), and the second factor is the deadline to perform it.

2.2. Dominance

We define the dominance of strategy A over strategy B, when under certain conditions, strategy A implies a better result than strategy B in terms of test results, i.e., more questions are correctly answered. To illustrate the situation, we consider a reading time ($r_i = r = 1$) and a re-reading time ($l_i = l = 1$) to be equivalent to one time slot, the response times for the hard question (p_h) and the easy question (p_e) equals to 7 and 2 time slots, respectively; and the test sequence given by S={h,e}. Figure 1 shows that in the case of a deadline $\overline{d}$ of four slots of time. On the one hand, if the RA strategy is used, no question will be answered because it will answer the hard question in the first instance, whose resolution time is longer than the deadline. On the other hand, if the RSA strategy is used, one question can be answered because the first question must have been skipped, and the easy question will have been read and answered. In this case, the RSA strategy dominates RA strategy. In a similar way, Figure 2 shows the case of a deadline $\overline{d}$ of ten slots, where the RA strategy will dominate with two questions answered over the RSA strategy, which will have only managed to answer a single question.

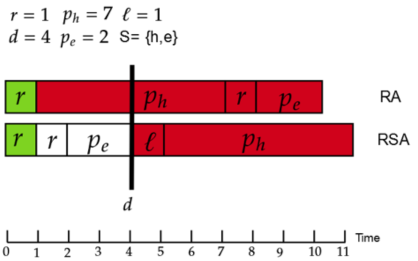


Figure 1. Example of dominance RSA over RA.

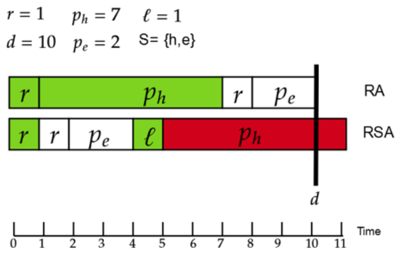


Figure 2. Example of dominance RA over RSA.

2.3. Mathematical formulation

The problem can be formulated as *integer linear programming*. We adopt a upper bound on the time to perform the test given by $K = \sum_{i=1}^{|P|} r_i + p_i + l_i$. Consequeltly, we discretize the time into K unit time slots. We then define three sets of binary decisión variables: $X_{i,t}$ which is equal to 1 if the reading of question i starts at time t and 0 otherwise; $Y_{i,t}$ which is equal to 1 if the answering of question i starts at time t and 0 otherwise; and $Z_{i,t}$ which is equal to 1 if the re-reading of question i starts at time t and 0 otherwise. The problem can be set as follows:

$$[Max] \sum_{i=1}^{|\mathcal{P}|} \sum_{t=1}^{\bar{d}-p_i+1} Y_{i,t} \qquad (1)$$

$$X_{i,t} = 0, \forall i \in \mathcal{P}, \forall t \in [K - r_i + 1, K] \qquad (2)$$

$$Y_{i,t} = 0, \forall i \in \mathcal{P}, \forall t \in [K - p_i + 1, K] \qquad (3)$$

$$Z_{i,t} = 0, \forall i \in \mathcal{P}, \forall t \in [K - \ell_i + 1, K] \qquad (4)$$

$$\sum_{t=1}^{K} X_{i,t} \leq 1, \quad \forall i \in \mathcal{P} \tag{5}$$

$$\sum_{t=1}^{K} Y_{i,t} \leq 1, \quad \forall i \in \mathcal{P} \tag{6}$$

$$\sum_{t=1}^{K} Z_{i,t} \leq 1, \quad \forall i \in \mathcal{P} \tag{7}$$

$$X_{i,t} = 1, \quad i = 1, \quad t = 1 \tag{8}$$

$$\sum_{i=1}^{|\mathcal{P}|} \left(\sum_{t=\max\{k-p_i,0\}}^{k-1} Y_{i,t} + \sum_{t=\max\{k-r_i,0\}}^{k-1} X_{i,t} + \sum_{t=\max\{k-\ell_i,0\}}^{k-1} Z_{i,t} \right) \leq 1, \quad K = 1 \tag{9}$$

$$\sum_{i=1}^{|\mathcal{P}|} \left(\sum_{t=\max\{k-p_i,0\}}^{k-1} Y_{i,t} + \sum_{t=\max\{k-r_i,0\}}^{k-1} X_{i,t} + \sum_{t=\max\{k-\ell_i,0\}}^{k-1} Z_{i,t} \right) \geq \sum_{i=1}^{|\mathcal{P}|} \left(\sum_{t=\max\{k-p_i,0\}}^{k-1} Y_{i,t} + \sum_{t=\max\{k-r_i,0\}}^{k-1} X_{i,t} + \sum_{t=\max\{k-\ell_i,0\}}^{k-1} Z_{i,t} \right), \quad \forall\, k \in K \tag{10}$$

$$X_{i+1,t} \leq \sum_{k=1}^{t} X_{i,k}, \forall\, i \in \mathcal{P} - 1, \forall\, t \in K \tag{11}$$

$$Z_{i,t} \leq \sum_{k=1}^{t} X_{i,k}, \forall\, i \in \mathcal{P}, \quad \forall\, t \in K \tag{12}$$

$$Z_{i,t} \leq Y_{i,t+\ell_i}, \quad \forall\, i \in \mathcal{P}, \quad \forall\, t \in K - \ell_i \tag{13}$$

$$1 - X_{i,t} \geq \sum_{k=1}^{t} Y_{i,k}, \forall\, i \in \mathcal{P}, \forall\, t \in K \tag{14}$$

$$1 - X_{i,t} \geq \sum_{k=1}^{t} X_{i+1,k}, \forall\, i \in \mathcal{P} - 1, \forall\, t \in K \tag{15}$$

$$\sum_{t=1}^{K} \left((t + r_i) \cdot X_{i,t} \right) \leq \bar{d}, \forall\, i \in \mathcal{P} \tag{16}$$

$$\sum_{t=1}^{K} \left((t + p_i) \cdot Y_{i,t} \right) \leq \bar{d}, \quad \forall\, i \in \mathcal{P} \tag{16}$$

$$\sum_{t=1}^{K} \left((t + \ell_i) \cdot Z_{i,t} \right) \leq \bar{d}, \quad \forall\, i \in \mathcal{P} \tag{18}$$

$$\sum_{t=1}^{K} t \cdot Y_{i,t} - \sum_{t=1}^{K} \left((t + r_i) \cdot X_{i,t} \right) \leq 1 + M \cdot \left(\sum_{t=1}^{K} Z_{i,t} \right), \quad \forall\, i \in \mathcal{P} \tag{19}$$

2.4. Implementation and computational experiments

The mathematical formulation described above was implemented computationally using Python 3.9.16. The Pyomo package was used for implementation and the Gurobi solver 10.0 was used for solving. For the computational experiments were performed on a 3-question test. To perform the analysis, the parameters were set to $r_i = \ell_i = 1$. For each test, a grid of values for the response times of each question will be made. This grid will have a lower limit of 1 time slot and an upper limit of 6 time slots. The value of the response time for difficult questions will be a function of the value of the response time for easy questions. e.g. If the answer time for the easy question is 1, the answer time for the difficult questions will be 2, 3, 4, 5 and 6 time slots. In turn, the deadline will be varied, for each of the cases described above, the deadline will be decreasing. In addition, the value $\overline{d}$ will be generated as an upper bound, this will have a value of 24 time slots, this will be calculated by using the following expression: $\overline{d} = \sum_{i=1}^{3} r_i + p_i + l_i$ where all the questions will be difficult. Therefore the number of instances to solve will be of $2^3 \cdot 5 \cdot 4 \cdot 3 \cdot 2 = 960$.

3. Results

In the generation of instances for their respective analysis, all possible test sequences were considered, that is, 2^3 possible combinations. These combinations are: {{h,h,e}; {e,e,e}; {h,e,e}; {h,h,e}; {e,e,h}; {e,h,h}; {e,h,e}; {h,e,h}}. Since it is assumed that the test-taker will know the number of questions and the sequence of the test, 5 instances will be discarded, since the RA strategy will dominate over RSA strategy regardless of the deadline. For example, if the test sequence is {e,e,e} there will be no difficult question to skip so the RSA strategy will not exist. On the other hand, if the sequence is {h,h,h} the use of the RSA strategy would not be logical, since knowing the sequence of the test, the use of this strategy would generate the additional cost of re-reading.

3.1. Sequential analysis {h,e,h}

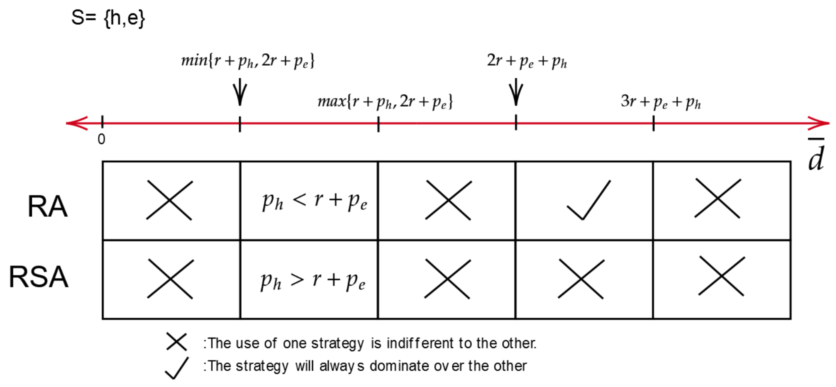
It was found that the use of the strategies studied is indifferent if the deadline is less than $\boldsymbol{min\{r + p_h, 2r + p_e\}}$ since the time to solve the test will not be enough to answer the first question, so the use of one or the other strategy will lead to the same result, not answering any question. On the other hand, we suppose the use of one strategy is indifferent to the other when the time is more than enough to perform the test in its entirety so that if the deadline is greater than $\boldsymbol{4r + p_e + 2p_h}$ no strategy will dominate over the other. Note that when facing the sequence {h,e}, one strategy will dominate over the other, according to the values of p_h and p_e, if the test deadline is between the values $\boldsymbol{min\{r + p_h, 2r + p_e\}}$ and $\boldsymbol{max\{r + p_h, 2r + p_e\}}$. Thus, if $\boldsymbol{p_h > r + p_e}$ the RSA strategy will dominate over RA strategy, otherwise, if $\boldsymbol{p_h < r + p_e}$ the RA strategy will dominate over RSA strategy. Finally, if the test-taker is facing the last question of the test, and his deadline is between the values $\boldsymbol{3 + p_e + 2p_h}$ and $\boldsymbol{4 + p_e + 2p_h}$, the dominant strategy will be the RA strategy.

3.2. Sequential analysis {e,h,e}

The result was that the use of the strategies studied is indifferent if the deadline is less than $\boldsymbol{min\{2r + p_e + p_h, 3r + 2p_e\}}$ since the test resolution time will be sufficient to answer the first question, that is, the easy question (facing an easy question does not require the use of a strategy) but it will not be sufficient to answer the next question. On the other hand, the use of one strategy will be indifferent to the other is when the time is more than enough to perform the test in its entirety, so if the deadline is greater than $\boldsymbol{4r + 2p_e + p_h}$ no strategy will dominate over the other. If the test deadline is between the values $\boldsymbol{min}\ \{2r + p_e + p_h, 3r + 2p_e\}$ and $\boldsymbol{max}\ \{2r + p_e + p_h, 3r + 2p_e\}$ one strategy will dominate over the other according to the values of p_h and p_e. If $p_h > r + p_e$ the RSA strategy will dominate over RA strategy, on the other hand if $p_h < r + p_e$ the RA strategy will dominate over RSA strategy. Finally, if the test-taker is facing the last question of the test, and its deadline is between the values $3 + 2p_e + p_h$ and $4 + 2p_e + p_h$ the dominant strategy will be the RA strategy.

3.3. Sequential analysis {h,e,e}

The result was that the use of the strategies studied is indifferent if the deadline is less than $\boldsymbol{min}\{r + p_h, 2r + p_e\}$ since the test resolution time will not be enough to answer the first question, so that the use of one or the other strategy will lead to the same result, not answering any question. In addtion, the use of one strategy will be indifferent to the other is when the time is more than enough to perform the test in its entirety, so if the deadline is greater than $4r + 2p_e + p_h$ no strategy will dominate over the other. If the test deadline is between the values $\boldsymbol{min}\{r + p_h, 2r + p_e\}$ and $\boldsymbol{max}\{r + p_h, 2r + p_e\}$ one strategy will dominate over the other according to the values of p_h and p_e. If $p_h > r + p_e$ the RSA strategy will dominate over RA strategy, on the other hand if $p_h < r + p_e$ the RA strategy will dominate over RSA strategy. Finally, if the test taker is facing the last question of the test, and its deadline is between the values $3 + 2p_e + p_h$ and $4 + 2p_e + p_h$ the dominant strategy will be the RA strategy. Note that in the tests performed similarities were found in the analysis, these similarities were found when the sequence (S) of the test is {h,e}. Figure 3 shows a graphical summary of what was described above.



bub

Figure 3. Summary table of dominance.

4. Conclusion and future works

Previous research has focused on practical use with the objective of finding strategies to perform different types of tests, while in the area of psychology, various strategies have been investigated with the objective of reducing stress and anxiety when facing a test. However, these studies have not been approached from the algorithms, mathematics, and optimization viewpoints. This study generated a mathematical scheduling model that generated a standardized test with the objective of applying the strategies found in the literature in order to find dominance properties. By performing several instances of a standardized three-question test, it was concluded that one strategy would dominate over the other depending on the parameters and design of the test. In particular, we observe that the use of one strategy over the other will be indifferent if the maximum time to perform the test is not enough to answer the first question or if the maximum time is more than enough to answer the whole test. Finally, we show that the RA strategy will always dominate over the RSA strategy when the last question of the test is asked. Studying these dominance properties can provide a guide for test-takers to tackle standardized tests and obtain an expected result. For future works, we leave open the problem with a sequence and number of unknown questions by the test-taker, or where each question could have weights or a probability of error, increasing the difficulty of the problem to be studied.

Acknowledgments

The authors are grateful for the support of the National Agency for Research and Development (ANID), FONDECYT Project N°1211640.

References

DEMRE. (2023, ENERO). Retrieved from https://demre.cl/la-prueba/pruebas-y-temarios/presentacion-pruebas-temarios-paes-invierno-p2024

Dufossé, F., Dürr, C., Nadal, N., Trystram, D., & Vásquez, Ó. C. (2022). Scheduling with a processing time oracle. *Applied Mathematical Modelling*, 701-720.

Dürr, C., Erlebach, T., Megow, N., & Meißner, J. (2020). An Adversarial Model for Scheduling with Testing. Algorithmica, 3630-3675.

Kesselman-Turkel, J., & Peterson, F. (1981). Test-Taking Strategies. Study smart series.

Retsef, L., Magnanti , T., & Shaposhnik, Y. (2018). Scheduling with Testing. Supplemental Materials.

Theobald, M., Breitwieser, J., & Brod, G. (2022). Test Anxiety Does Not Predict Exam Performance When Knowledge Is Controlled For: Strong Evidence Against the Interference Hypothesis of Test Anxiety. Psycological Science, 10.

Viet, V. V. (2022). Learning strategy and its effect on academic achievement. *Kasetsart Journal of Social Sciences*, 9.

Wang, T., & Zhang, J. (2006). Optimal partitioning of testing time: Theoretical properties and practical implications. *Psychometrika*, 15.