

Questões calculadas Moodle: evitam a cola?

Calculated questions in Moodle: do they prevent cheating?

Preguntas calculadas en Moodle: ¿evitan el copiar y pegar?

Recebido: 20/06/2023 | Revisado: 30/06/2023 | Aceitado: 02/07/2023 | Publicado: 08/07/2023

Alberto Bastos do Canto Filho

ORCID: <https://orcid.org/0000-0003-0822-3797>
Universidade Federal do Rio Grande do Sul, Brasil
E-mail: alberto.canto@ufrgs.br

Resumo

Durante o Ensino Remoto Emergencial foram suspensas todas as avaliações presenciais trazendo grande desconforto e questionamentos sobre a autenticidade das respostas dadas às provas elaboradas a distância sem a existência de uma supervisão para garantir que os estudantes não copiassem as respostas dos colegas (cola). Uma das opções que foram exploradas para minimizar o risco de cola foi desenvolver questionários com respostas numéricas diferentes para cada estudante, explorando a funcionalidade Questões Calculadas, disponibilizada no ambiente virtual de aprendizagem Moodle. Como saber se esta estratégia foi eficaz? Em que proporção podemos confiar que a exigência de uma resposta numérica diferente para cada estudante aumentará a probabilidade de que a resposta correta tenha sido efetivamente calculada pelo respondente? Neste artigo foram analisados 12 semestres letivos (desde 2016 até 2021) ministrados para Cursos de Engenharia utilizando questionários semanais assíncronos compostos por Questões Calculadas Moodle. Utilizando-se a informação de “tempo para responder” foram obtidas curvas típicas de distribuição, considerando a primeira vez que as questões são apresentadas. Observou-se que, após algum tempo, os tempos médios e desvios padrões se reduzem, indicando que os estudantes estão obtendo as respostas sem aplicar o tempo que seria necessário para a realização de cálculos complexos exigidos para obter a resposta. O objetivo deste trabalho é apresentar uma metodologia capaz de, a partir de relatórios padrão do Moodle, coletar indicadores sobre a eficácia do uso de “Questões Calculadas” para evitar a cola.

Palavras-chave: Ensino; Avaliação formativa; Aprendizagem baseada em problemas; Avaliação à distância; Credibilidade das provas.

Abstract

During the Emergency Remote Teaching, all in-person assessments were suspended, causing great discomfort and raising questions about the authenticity of the answers given in remotely administered exams without the presence of supervision to ensure that students did not copy answers from their peers (cheating). One of the options explored to minimize the risk of cheating was to develop quizzes with different numerical answers for each student, using the Calculated Questions feature provided in the Moodle virtual learning environment. How can we determine if this strategy was effective? To what extent can we trust that the requirement for a different numerical answer for each student will increase the likelihood that the correct answer was actually calculated by the respondent? In this article, 12 academic semesters (from 2016 to 2021) taught in Engineering courses were analyzed, using asynchronous weekly quizzes composed of Moodle Calculated Questions. By using the "time to respond" information, typical distribution curves were obtained, considering the first time the questions were presented. It was observed that, over time, the average response times and standard deviations decrease, indicating that students are obtaining the answers without applying the time that would be necessary for complex calculations required to obtain the answer. The objective of this study is to present a methodology capable of collecting indicators on the effectiveness of using "Calculated Questions" to prevent cheating, based on standard Moodle reports.

Keywords: Teaching; Formative assessment; Problem-based learning; Exams taken at home; Exams' credibility.

Resumen

Durante la Enseñanza Remota de Emergencia, se suspendieron todas las evaluaciones presenciales, lo que generó gran malestar y planteó interrogantes sobre la autenticidad de las respuestas dadas en los exámenes elaborados a distancia sin la supervisión necesaria para garantizar que los estudiantes no copiaran las respuestas de sus compañeros (hacer trampa). Una de las opciones exploradas para minimizar el riesgo de hacer trampa fue desarrollar cuestionarios con respuestas numéricas diferentes para cada estudiante, aprovechando la funcionalidad de las Preguntas Calculadas disponibles en el entorno virtual de aprendizaje Moodle. ¿Cómo saber si esta estrategia fue eficaz? ¿En qué medida podemos confiar en que la exigencia de una respuesta numérica diferente para cada estudiante aumentará la probabilidad de que la respuesta correcta haya sido efectivamente calculada por el encuestado? En este artículo se analizaron 12 semestres académicos (desde 2016 hasta 2021) impartidos en cursos de Ingeniería utilizando cuestionarios semanales

asincrónicos compuestos por Preguntas Calculadas de Moodle. Utilizando la información de "tiempo de respuesta", se obtuvieron curvas típicas de distribución considerando la primera vez que se presentaron las preguntas. Se observó que, con el tiempo, los tiempos promedio y las desviaciones estándar disminuyen, lo que indica que los estudiantes están obteniendo las respuestas sin aplicar el tiempo necesario para realizar cálculos complejos requeridos para obtener la respuesta. El objetivo de este trabajo es presentar una metodología capaz de recopilar indicadores sobre la eficacia del uso de las "Preguntas Calculadas" para evitar hacer trampa, basada en los informes estándar de Moodle.

Palabras clave: Enseñanza; Evaluación formativa; Aprendizaje basado en problemas; Evaluación a distancia; Credibilidad de los exámenes.

1. Introdução

O uso de questionários no Ensino a Distância pode ser adotado considerando uma perspectiva pedagógica ou de avaliação. Considerando a perspectiva pedagógica, a possibilidade de que um estudante copie os resultados obtidos pelos colegas não é necessariamente um aspecto negativo, desde que seja explorada adequadamente para fins de aprendizagem (avaliação formativa) (de Souza & Innocenti, 2022) ou motivacionais (Canto et al., 2012; Pacheco et al., 2022).

No entanto, quando o objetivo de um questionário é avaliar se o estudante alcançou um determinado objetivo educacional (avaliação somativa) (Belhot, 2010; Bloom & Krathwohl, 1956; Krathwohl, 2002; Tuchtenhagen et al., 2022), a prática da cola invalida o resultado da avaliação, pois este não estará expressando adequadamente a aprendizagem de cada estudante.

A premissa de que é impossível evitar a cola em provas realizadas a distância levou o Ministério da Educação (MEC) a exigir que os cursos a distância realizem exames presenciais, nos quais a prova é realizada sob uma supervisão que garanta a autenticidade das respostas (da Fonseca et al., n.d.; Presidência da República, 2017).

A pandemia COVID-19 trouxe consigo a necessidade de um isolamento social e o *Ensino Remoto Emergencial* (ERE) que, diferentemente do *Ensino a Distância* (EaD), adotou como padrão a realização de 100% das avaliações a distância. Isto é, dada a natureza emergencial do período de pandemia, relevou-se a um segundo plano a premissa de que *avaliações a distância podem ter seu resultado comprometido devido à possibilidade de cola*.

A vulnerabilidade do sistema de avaliação a distância foi motivo de desconforto para muitos docentes (da Silva & Afonso, 2021), levando-os a explorar diferentes métodos de avaliação (Herrera & Angeletti, 2021; mac Gaul et al., 2021), sempre tendo em mente a impossibilidade de implementar sistemas de fiscalização eletrônica (*proctoring*) (Nigam et al., 2021) pois muitos dos estudantes sequer dispunham de uma webcam.

Para minimizar o risco de cola durante o ERE, as *Questões Calculadas Moodle* (Moodle, 2022a) foram amplamente exploradas para substituir provas presenciais nos cursos de ciências exatas da Universidade Federal do Rio Grande do Sul (UFRGS). Estas questões exigem respostas diferentes para cada aluno o que, hipoteticamente, impossibilitaria que os estudantes copiassem as respostas dos colegas.

Infelizmente esta expectativa não se concretizou como desejado, pois os estudantes desenvolveram métodos para obter a resposta para sua questão utilizando memoriais de cálculo desenvolvido pelos colegas e disponibilizados na forma de planilhas eletrônicas ou fórmulas de com a resolução simbólica.

As questões calculadas possuem um *ciclo de vida*, que começa com a sua criação. Quando utilizadas pela primeira vez numa avaliação, estas questões possuem o poder de diferenciar os estudantes em função de seu aprendizado. Imediatamente após a obtenção da resposta correta passa a existir a possibilidade de disseminação do gabarito (memorial de cálculo) entre os estudantes.

Este artigo investiga alternativas capazes de mensurar o ciclo de vida das questões calculadas. As seguintes questões são investigadas:

- É possível utilizar os dados disponibilizados no Moodle para verificar se as questões calculadas perderam seu poder de discriminação?
- Como?

O artigo baseia-se em dados coletados ao longo de seis anos durante os quais foram utilizados questionários com *Questões Calculadas Moodle* em cursos de Engenharia da UFRGS. Seu objetivo é apresentar uma metodologia capaz de, a partir de relatórios padrão do Moodle, coletar indicativos sobre a eficácia do uso de “Questões Calculadas” para evitar a cola.

2. Metodologia

A metodologia apresentada nesta seção foi desenvolvida pelo autor inspirada em técnicas de manutenção preditiva, que utilizam métodos estatísticos para comparar a evolução ao longo do tempo de medidas realizadas em plantas industriais. Segundo as classificações propostas por Pereira et al. (2018), o método utilizado pode ser caracterizado como um método indutivo, tendo em vista que cria uma generalização a partir de uma situação particular, comparativo, pois compara determinados indicadores de um semestre com os semestres subsequentes e estatístico, pois utiliza de cálculos estatísticos.

Foram utilizados dados de problemas pertencentes a um banco de problemas, que vem sendo utilizado em diversas turmas/disciplinas de Cursos de Engenharia da Universidade Federal do Rio Grande do Sul (UFRGS). Estes problemas são implementados na forma de *Questões Calculadas Moodle* agrupadas em questionários utilizados por disciplinas que trabalhem com uma metodologia de *Aprendizagem Baseada em Problemas*. Foram analisados dados históricos coletados desde o primeiro semestre letivo de 2016, quando este banco de problemas começou a ser criado. Para garantir a possibilidade de análise comparativa, foi selecionada uma questão que permaneceu inalterada ao longo do período analisado.

Estes dados históricos foram utilizados com o objetivo de testar e especificar um conjunto de indicadores que, a partir de dados disponibilizados em relatórios padrão do Moodle, permitam identificar situações em as *Questões Calculadas Moodle* perdem o poder de discriminar se a resposta correta foi obtida a partir de cálculos efetivamente realizados pelo estudante ou foi obtida a partir de uma planilha de cálculo disponibilizada por um colega. A elaboração destes indicadores parte das seguintes premissas:

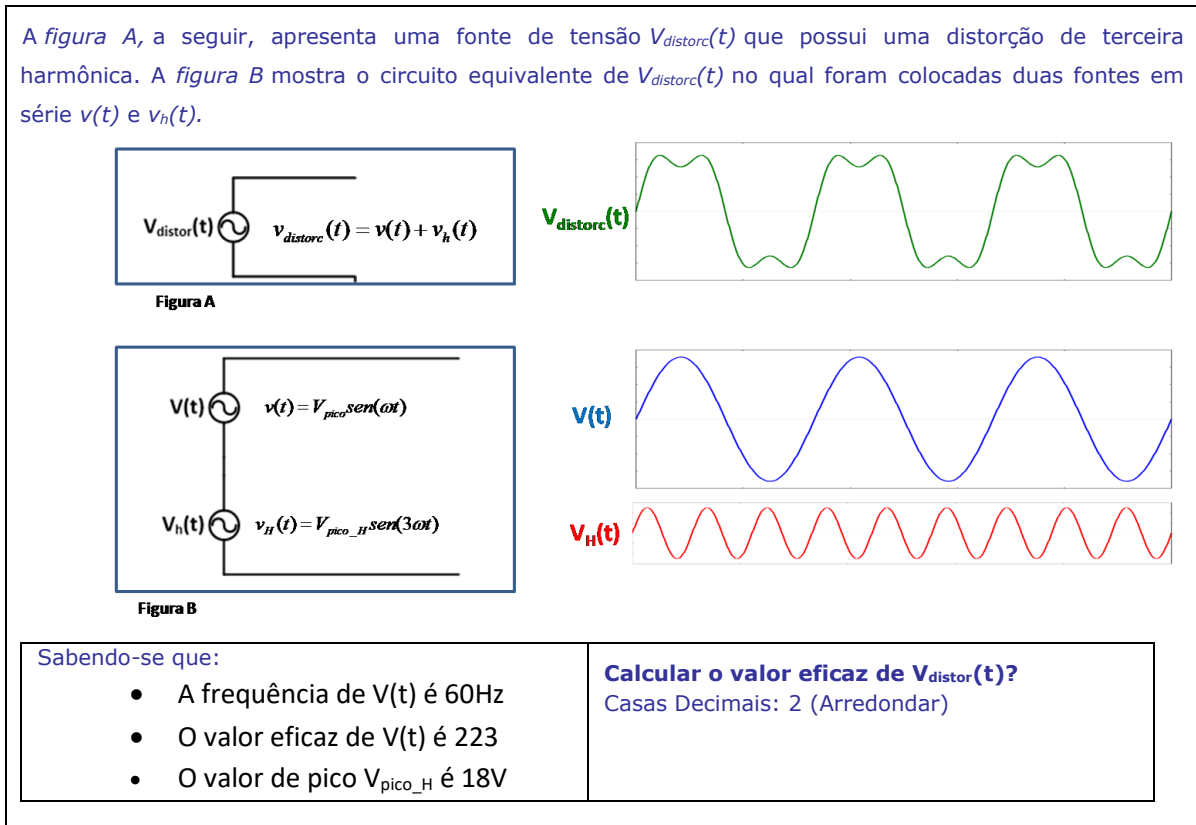
- O tempo necessário para obter a resposta correta de uma questão calculada é substancialmente menor quando se utiliza uma planilha de cálculo com fórmulas desenvolvidas por outros.
- O número de estudantes que acerta a questão cresce à medida que os recursos para cálculo da resposta correta sejam disseminados entre os estudantes.
- Uma questão inédita terá um percentual de tentativas corretas inferior a uma questão cujo gabarito já foi elaborado e disponibilizado.

2.1 Análise dos dados históricos

A investigação realizada considerou 12 semestres letivos (2016-1 a 2021-2) de disciplinas/turmas ministradas para cursos de Engenharia que utilizam metodologia de *Aprendizagem Baseada em Problemas*. O sistema de avaliação formativa foi implementado com *Questões Calculadas Moodle* (Moodle, 2022b) construídas com *Datasets* contendo 100 registros. Com o objetivo de motivar os estudantes a resolver os problemas propostos (Canto et al., 2012), somente são aprovados estudantes que obtenham um grau superior ou igual a 60% nestes questionários. Os estudantes têm um prazo de uma semana para responder ao questionário, e um número ilimitado de tentativas, valendo a nota mais alta.

Para realização desta investigação, analisou-se a segunda questão do questionário sobre “Valor Médio e Valor Eficaz”, mostrada na Figura 1. Esta questão foi escolhida pelo fato ter sido reutilizada sem alterações ao longo dos doze semestres analisados.

Figura 1- Questão Analisada na Investigação.



Fonte: Elaborada pelos autores.

Observe que a solução do problema apresentado na Figura 1 exige cálculos elaborados e, conseqüentemente, necessita de um tempo de desenvolvimento dos cálculos.

Os dados de cada um dos semestres analisados neste trabalho foram extraídos do Moodle, utilizando a funcionalidade de *download* do *Relatório detalhado de tentativas por aluno*. O processamento destas informações foi realizado utilizando um aplicativo de suporte desenvolvido especificamente para este fim. Este aplicativo será descrito na subseção a seguir.

2.2 MCQAS – Moodle Calculated Question Analysis System

O aplicativo *MCQAS – Moodle Calculated Question Analysis System* foi desenvolvido com o objetivo específico de analisar a perda de discriminação de questões calculadas. Programado na linguagem VBA (*Visual Basic for Application*), o aplicativo gera um conjunto de indicadores, a partir do *Relatório detalhado de tentativas por aluno*.

Os processamentos e indicadores gerados por este aplicativo são apresentados nas seções a seguir.

Classificação das tentativas quanto à correção e ao tempo

Neste procedimento, cada uma das tentativas realizadas pelos estudantes é classificada em uma das cinco categorias apresentadas na Tabela 1, abaixo.

Tabela 1 - Classificação das tentativas.

Categoria	Característica da tentativa
<i>InvRespNull</i>	Inválida por não responder; o estudante enviou o questionário sem responder à questão analisada.
<i>CorrectCoherentTime</i>	Estudante apresentou a resposta correta; o tempo total conectado ao questionário é maior ou igual ao tempo mínimo previsto para realizar os cálculos necessários. Neste estudo o parâmetro utilizado como tempo mínimo foi de 15 min.
<i>IncorrectCoherentTime</i>	Estudante apresentou a resposta incorreta; o tempo total conectado ao questionário é maior do que tempo mínimo previsto para realizar os cálculos.
<i>CorrectIncoherentTime</i>	Estudante apresentou a resposta correta; o tempo total conectado é inferior ao tempo mínimo previsto para realizar os cálculos necessários.
<i>IncorrectInsufficientTime</i>	Estudante apresentou a resposta incorreta; o tempo total conectado ao questionário é inferior ao tempo mínimo previsto para realizar os cálculos.

Fonte: Elaborado pelos autores.

A classificação apresentada na Tabela 1, considera dois critérios:

- correção da resposta;
- tempo de conexão.

As tentativas são também classificadas como *LongTry*: um campo booleano que assume o valor *verdadeiro* quando o *tempo dedicado a responder à questão* foi maior que o parâmetro *tempo máximo*, especificado com o valor de quatro horas para este estudo. Assume-se que, nas tentativas *LongTry*, o *Tempo Utilizado* não corresponde ao *Tempo de Trabalho*, isto é, houve uma interrupção da resolução da questão e posterior retomada, gerando um valor irreal sobre o tempo que o estudante efetivamente se dedicou a resolver o questionário.

Indicadores ‘por aluno’ e ‘por semestre’

Após a extração dos dados brutos do Moodle (download do *Relatório detalhado de tentativas por aluno*), ocorre a primeira etapa de processamento, na qual o aplicativo MCQAS calcula um conjunto de indicadores ‘por aluno’ e ‘por semestre’.

Estes indicadores são as seguintes totalizações para cada aluno/semestre letivo:

- **Correct Coerent Time:** informa o número de tentativas com resposta correta e tempo coerente com a dificuldade do questionário.
- **Correct Incoherent Time:** informa o número de tentativas com resposta correta e tempo inferior ao tempo mínimo previsto para responder ao questionário.
- **Incorrect Coherent Time:** informa o número de tentativas com resposta incorreta e tempo coerente com a dificuldade da questão.
- **Inv Resp Null:** informa o número de tentativas para as quais não foi informado um valor de resposta para a questão analisada.
- **Incorrect Insufficient Time:** informa o número de tentativas com resposta incorreta e tempo inferior ao tempo mínimo previsto para responder ao questionário;
- **OK:** informa se o estudante apresentou pelo menos uma tentativa respondendo corretamente à questão analisada.
- **Trys Before OK:** informa o número de tentativas com resposta incorreta antes de enviar a primeira tentativa com resposta correta.
- **Time To OK:** informa o tempo total despendido em tentativas até acertar (inclui tempo da primeira tentativa correta).
- **Time Before OK:** informa o tempo total despendido em tentativas com resposta incorretas antes de acertar.

A Tabela 2, abaixo, mostra um exemplo destes indicadores.

Tabela 2 - Indicadores *por aluno e por semestre*.

Semester	Email	Correct	Correct	Incorrect	Inv Resp	Incorrect	trys		Time Before OK
		Coherent Time	Incoherent Time	Coherent Time		Insufficient Time	OK	Before OK	
161	xxx1				2		0	0	00 00:00:00
161	xxx2	2	4	1			1	1	00 00:42:19
161	xxx3	3		1			1	1	00 02:22:32
162	xxx4		1			1	1	1	00 00:05:56
162	xxx5		1	1			1	1	00 01:11:12
162	xxx6				3		0	0	00 00:00:00
162	xxx7				2		0	0	00 00:00:00

Fonte: Elaborado pelos autores.

Observe na Tabela 2.

- A primeira linha apresenta os indicadores do aluno identificado pelo e-mail *xxx1* no primeiro semestre de 2016 (Semester=161): foram feitas duas tentativas de responder ao questionário; em nenhuma delas foi fornecida a resposta para a questão analisada (Inv Resp Null=2);
- A segunda linha apresenta os indicadores do estudante identificado pelo e-mail *xxx2* no primeiro semestre de 2016 (Semester=161): este estudante realizou: a) duas tentativas com resposta correta e tempo coerente (Correct Coherent Time=2); b) quatro tentativas com resposta correta e tempo muito baixo (Incoherent Time=4); c) uma tentativa com resposta incorreta e tempo coerente com a dificuldade do questionário (Incorrect Coherent Time=1). O tempo total despendido até acertar (Time To OK) a questão foi de 42 minutos e 19 segundos; despendeu 26 minutos e 16 segundos em tentativas que errou a resposta dada à questão analisada (Time To OK).

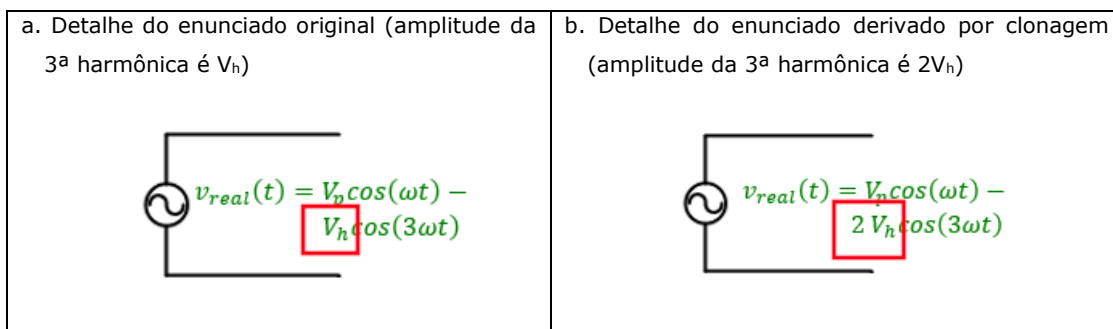
Cabe esclarecer que esta questão faz parte de um conjunto de questões que compõem um questionário; mesmo tendo acertado esta questão numa determinada tentativa, eventualmente o estudante faz uma nova tentativa buscando acertar uma outra questão do questionário.

A partir dos indicadores ‘por aluno’ e ‘por semestre’ o aplicativo MCQAS calcula *Tempo médio até Responder* como a mediana do indicador *Time To OK*, considerando todas as tentativas válidas de um determinado semestre.

2.3 Questões aleatórias e inéditas

Conforme será apresentado na seção 3 - *Resultados encontrados*, constatou-se que, com o passar do tempo houve uma redução do *Tempo para Responder*, o que, segundo as premissas desta pesquisa, é um indício de que os estudantes estão colando. A partir do primeiro semestre de 2022 a questão analisada foi substituída por uma *Questão Aleatória (Moodle Ramdom Question, 2022)*. Esta funcionalidade permite que um questionário apresente uma questão sorteada no momento em que o estudante inicia a tentativa. Criou-se a possibilidade de que fosse sorteada aleatoriamente a questão original ou uma outra questão que foi criada num processo de *Derivação por Clonagem*, que gerou uma questão idêntica à original, exceto por uma sutil alteração no enunciado que implica a realização de cálculos diferentes para obtenção do resultado correto. A Figura 2 mostra o único detalhe alterado na questão criada desta forma.

Figura 2 - Diferença entre a questão original e a questão derivada por clonagem.



Fonte: Elaborado pelos autores.

Observe que a resolução simbólica da questão original (Figura 2a) é diferente da resolução simbólica questão modificada (Figura 2b). Com isto o uso de uma planilha eletrônica desenvolvida para a questão original resultará numa resposta incorreta para a questão modificada.

Os estudantes que foram aleatoriamente sorteados para responder à questão original foram considerados pertencentes ao *Grupo A*; os estudantes sorteados para responder a variante derivada por clonagem foram considerados pertencentes ao *Grupo B*.

A análise comparativa dos resultados obtidos por estes dois grupos é apresentada na subseção 3.3.

3. Resultados e Discussão

A investigação realizada considerou 12 semestres letivos (2016-1 a 2021-2) de uma disciplina ministrada para estudantes de engenharia. Durante estes doze semestres, a questão mostrada na Figura 1 acima, foi reutilizada em um questionário desenvolvido num contexto de *Avaliação Formativa e Aprendizagem Baseada em Problemas*.

A Tabela 3 abaixo, mostra um sumário contendo o número de tentativas realizadas neste período.

Tabela 3 - Tentativas de responder o questionário.

Tentativas	Número	%
Processadas:	1530	56%
Incorreta	575	21%
Resposta Correta	955	35%
Descartadas:	1188	44%
Excesso de tempo	388	14%
Sem Resposta	800	30%
Total	2718	100%

Fonte: Elaborado pelos autores.

Observe na Tabela 3 que 44% das tentativas realizadas foram descartadas na análise dos dados realizada neste estudo. Na subseção 3.1, a seguir, será apresentada a razão pela qual estas tentativas foram descartadas.

3.1 Tentativas descartadas

A premissa de que “os estudantes que efetivamente resolvem a questão necessitam um tempo superior àqueles que respondem à questão utilizando um gabarito elaborado por outros” somente pode ser mensurada se houver uma forma de obter

o tempo que cada estudantes efetivamente se dedicou a responder à questão. Esta informação não está disponível nos relatórios disponibilizados pelo Moodle: a informação disponível é o “*Tempo para completar o Questionário*”, calculado como o tempo total transcorrido entre o início da tentativa e o seu envio.

Constatou-se que, como o período de disponibilização do questionário foi de sete dias, alguns estudantes optaram por abrir o questionário no primeiro dia e somente enviá-lo alguns dias depois. Nestes casos a “*Tempo para completar o Questionário*” não representa o tempo que estudante efetivamente se dedicou a resolver ao questionário, razão pela qual optou-se por desconsiderar *por excesso de tempo* as tentativas com duração superior a quatro horas (tentativas classificadas como *LongTry*). Isto representou a exclusão de 388 tentativas (14%).

Foram também excluídas todas as tentativas *sem resposta* (800 tentativas; 30%), correspondentes aos casos em que o estudante enviou a tentativa de responder ao questionário sem informar uma resposta para a questão analisada.

Após as exclusões *por excesso de tempo* e das tentativas *sem resposta*, restaram 1530 tentativas (56%), que foram utilizadas para calcular o “*Tempo até Acertar*”.

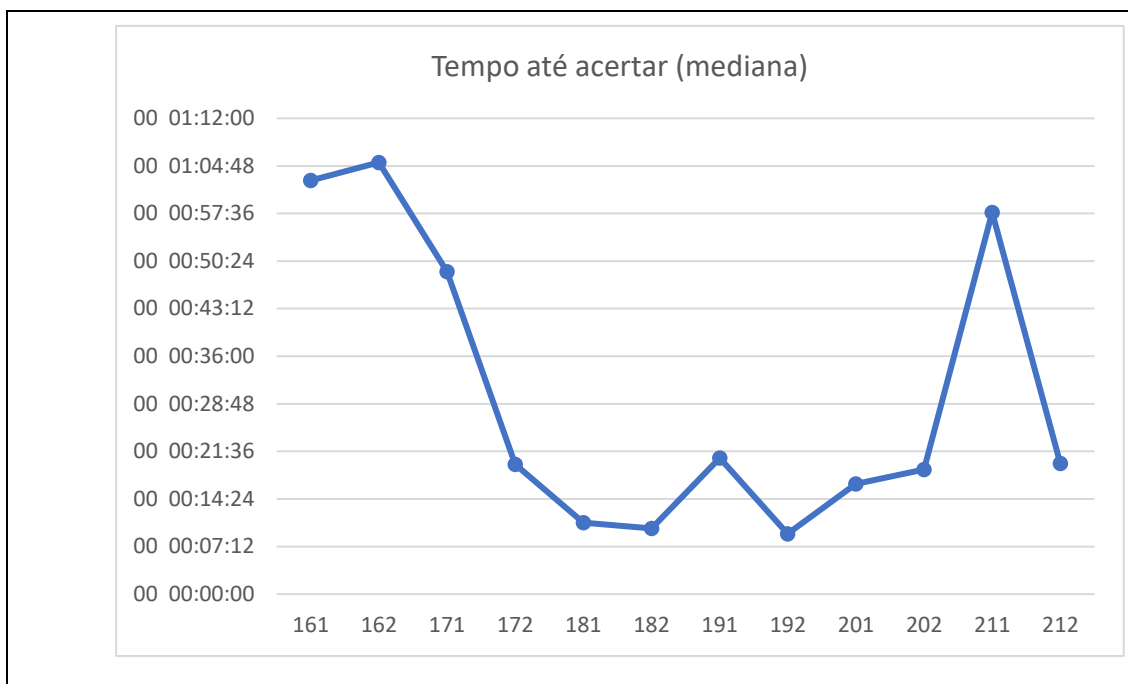
3.2 Tempo até acertar

Nas disciplinas/turmas analisadas neste artigo, os questionários Moodle foram configurados para aceitar um número ilimitado de tentativas durante o período de disponibilização do questionário (uma semana), valendo a nota mais alta obtida. Neste tipo de configuração o cálculo do tempo dedicado à resolução da questão deverá considerar três cenários:

1. O estudante faz várias tentativas; em todas erra a questão: neste cenário não será possível aferir o tempo dedicado até acertar a questão.
2. O estudante faz várias tentativas, até acertar a questão: neste cenário o tempo dedicado será o somatório dos tempos das tentativas, incluindo a primeira tentativa que apresentou a resposta correta.
3. O estudante faz várias tentativas erradas até acertar; depois de acertar realiza novas tentativas porque ainda não acertou as outras questões do questionário: neste cenário o tempo para acertar deve desconsiderar o tempo dedicado às tentativas posteriores ao primeiro acerto.

A Figura 3, abaixo, mostra a evolução do indicador *Tempo até acertar* desde o primeiro semestre de 2016 (semestre 161) até o segundo semestre de 2021 (semestre 212). O eixo horizontal mostra o semestre e o eixo vertical mostra a mediana do *Tempo até acertar*. Optou-se por utilizar a *mediana* (e não a *média*) porque a informação de *Tempo para acertar* não passou nos testes de verificação de distribuição segundo uma curva normal.

Figura 3 - Tempo até acertar - evolução semestral.



Fonte: Elaborado pelos autores.

Observe na Figura 3, acima que, no primeiro semestre que utilizou este questionário (semestre 161), a mediana foi de uma hora e dois minutos. Isto significa que metade dos alunos enviou o questionário com a resposta correta num tempo inferior a 62 minutos.

Nota-se uma tendência de queda deste indicador até o segundo semestre de 2018 (semestre 182), quando a metade dos alunos matriculados conseguiu enviar a resposta correta em um tempo inferior a 9 minutos.

Considerando a complexidade da questão analisada, pode-se afirmar que esta redução do tempo até acertar é uma evidência de que o seu gabarito está amplamente difundido entre os estudantes e que, portanto, esta questão perdeu a sua capacidade de avaliar se o estudante a acertou realmente ou se utilizou um memorial de cálculo desenvolvido por colegas para obter a resposta correta.

Observe ainda, na Figura 3:

- No primeiro semestre de 2019 (semestre 191) houve um pico para o qual não se identificou uma justificativa. Trata-se de uma evidência de que o *tempo até acertar* é influenciado por diversas outras variáveis, tais como feriados, outras provas na mesma semana, perfil dos participantes da turma ou outros fatores a serem investigados em trabalhos futuros.
- No primeiro semestre de 2020 (semestre 201) iniciou-se o *Ensino Remoto Emergencial*: as aulas e avaliações que eram presenciais até o semestre anterior foram substituídas por videoaulas e questionários Moodle. A partir deste semestre há indicativos de estabelecimento de um novo patamar de mediana tendendo a 20 minutos.
- No primeiro semestre de 2021 (semestre 211) ocorre um pico no tempo médio por tentativa (mediana de 58 minutos). Neste semestre uma das outras questões do questionário foi substituída (a questão analisada neste estudo foi mantida inalterada). O tempo para enviar o questionário foi maior porque o gabarito desta nova questão ainda não havia sido difundido amplamente.

- No segundo semestre de 2021 (semestre 212) o tempo até acertar reduziu-se novamente, indicando que o tempo para disseminação dos gabaritos passou a ser inferior a um semestre. Isto é, o tempo necessário para disseminação dos memoriais de cálculo reduziu-se desde 2016, quando se começou a utilizar questões calculadas, até o presente momento, em que os alunos já aprenderam a produzir e disseminar planilhas de cálculo que permitem obter as respostas corretas.

Com base na constatação de que o tempo para disseminar os memoriais de cálculo das Questões Calculadas é inferior à um semestre letivo, a partir do primeiro semestre letivo de 2022 o questionário foi alterado, substituindo a questão que vinha sendo reutilizada há doze semestres sem alterações por uma *Questão Aleatória* sorteada num universo que contempla duas possibilidades: a questão original (mostrada na Figura 1 e uma questão derivada por clonagem (descrita na subseção 2.3 acima).

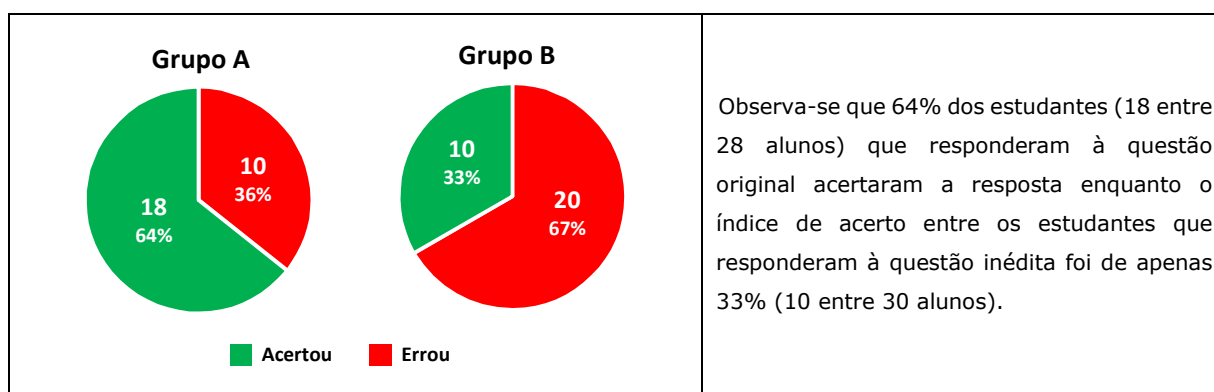
Na subseção a seguir é apresentado o impacto desta alteração.

3.3 Questões inéditas/aleatórias

A Figura 4, abaixo, mostra o efeito da introdução de questões aleatórias derivadas por clonagem. Conforme descrito na subseção 2.3, acima, o Grupo A é composto por estudantes que responderam ao questionário que contém uma questão reutilizada há 12 semestres e o Grupo B é composto por estudantes que responderam a uma questão muito parecida com a original, contendo apenas uma sutil diferença que inviabiliza a utilização de planilhas desenvolvidas para obter a resposta correta para a versão original.

Entre os estudantes que acertaram a questão no Grupo B (10 estudantes que responderam à versão inédita da questão), foi realizada uma análise das tentativas anteriores ao acerto. Constatou-se que 60% destes (6 estudantes) fizeram uma tentativa de responder esta questão apresentando os resultados que seriam obtidos utilizando as expressões de cálculo da questão respondida pelo Grupo A, evidenciando a prática de utilização de recursos externos para solucionar as questões calculadas.

Figura 4 - Percentuais de tentativas corretas.



Fonte: Elaborado pelos autores.

4. Considerações Finais

No ensino formal, a existência de prazos definidos para a realização de tarefas frequentemente entra em choque com o sistema de prioridades dos estudantes, que nem sempre dispõem do tempo necessário para a efetiva aprendizagem. Nestes casos, o receio de ser reprovado o leva a perceber o sistema de avaliação como um fator de motivação extrínseca (Gagné & Deci, 2005; Vansteenkiste et al., 2006) o que, ocasionalmente, o leva a apresentar os resultados obtidos por outros colegas para obter o grau necessário para aprovação. Além de eticamente questionável, a cola pode afetar a credibilidade dos professores e instituições, que terminam certificando o aprendizado de estudantes que não aprenderam efetivamente. Também é danoso quando se deseja

realizar qualquer avaliação do processo de aprendizagem (Molon et al., 2020), pois estes estudos estarão fundamentados em dados irreais e, portanto, apresentarão resultados incorretos.

Neste artigo foram apresentados alguns indicadores desenvolvidos com o objetivo de, com base em dados disponibilizados nos relatórios padrão do *Moodle*, verificar se as respostas dadas às questões calculadas foram obtidas a partir de cálculos efetuados pelo próprio estudante ou obtidas a partir de planilhas de cálculo desenvolvidas por colegas. Os dois indicadores explorados (*Tempo médio para acertar a questão* e *Percentual de tentativas incorretas*) mostraram-se parcialmente efetivos: quando a análise da evolução temporal destes indicadores indica uma tendência de queda, pode-se presumir uma provável disseminação dos gabaritos (planilhas de cálculo). No entanto, como o relatório padrão utilizado para cálculo destes indicadores disponibiliza apenas o *Tempo da Tentativa*, o uso deste valor como *Tempo Médio para Resolver a Questão* pode gerar incorreções devido à:

- existência de *questões aleatórias* no questionário;
- alteração do questionário (mudará o tempo das tentativas, mesmo que a questão analisada seja mantida inalterada);
- alterações nas configurações do questionário (prazo de entrega, número de tentativas, feedbacks, etc.).

Além dos aspectos citados acima os indicadores propostos podem também ser afetados por outros fatores intrinsecamente relacionados ao perfil dos estudantes:

- padrão ético dos integrantes da turma;
- motivação intrínseca pela aprendizagem;
- existência de outras demandas agendadas ou feriados em data próxima ao período de realização do questionário;
- forma de executar o questionário (de uma única vez, fazer uma questão por dia, etc.);
- docente responsável pela disciplina;
- melhoria das aulas ou dos métodos de ensino;
- modificações no sistema de avaliação da disciplina (por exemplo, introdução ou exclusão de avaliações somativas supervisionadas);
- outras variáveis que possam afetar o interesse, a aprendizagem ou forma como os estudantes respondem aos questionários.

Resumindo, os indicadores apresentados neste artigo dependem de diversas variáveis que, se alteradas, poderão comprometer o seu uso como indicador de que o estudante colou.

Existem alternativas que podem ser exploradas para reduzir a possibilidade de que os estudantes utilizem os resultados obtidos pelos colegas. Entre elas pode-se citar:

1. Deixar de atribuir grau aos problemas utilizados na avaliação formativa. Neste caso os questionários funcionariam como ‘listas de exercícios’ e os estudantes que copiam os resultados obtidos pelos colegas apenas para obter o grau necessário para aprovação não teriam motivo para despender seu tempo com esta atividade. Infelizmente são justamente estes estudantes que mais necessitam uma motivação adicional para que se dediquem a *Aprendizagem Baseada em Problemas*.
2. Utilizar questões inéditas. Esta alternativa dificulta a *cola*, mas representa um esforço docente relevante para produção contínua de questões inéditas.
3. As seguintes alternativas dificultam a *cola*, mas também podem dificultar a aprendizagem dos estudantes que estão motivados para empenhar um esforço legítimo de aprendizagem:

- Redução do tempo que o questionário fica disponibilizado;
 - Redução do número de tentativas;
 - Não mostrar o gabarito ao final de cada tentativa.
4. Utilizar questões aleatórias. Embora não tão eficaz como o uso de questões inéditas, o uso de questões aleatórias dificulta a obtenção da planilha de cálculo dos resultados, pois cada estudante necessitará realizar cálculos diferentes. O inconveniente desta alternativa é a necessidade de um *Banco de Questões* com vários problemas para cada questão

As soluções 2 e 4 citadas acima necessitam a produção de um grande número de questões calculadas e, conseqüentemente, um esforço docente para a produção de questões. Este esforço pode ser substancialmente reduzido se forem utilizados técnicas e recursos para produção sistemática de questões (Canto Filho, 2022; do Canto Filho, n.d.).

Finalmente, relata-se que, ao longo da realização deste estudo, constatou-se empiricamente um resultado positivo inesperado: os estudantes aprenderam autonomamente a utilizar planilhas eletrônicas para realização e documentação de memoriais de cálculo que pudessem reutilizar ou repassar aos colegas. O aspecto negativo desta nova competência adquirida autonomamente é o encurtamento do ciclo de vida das questões calculadas: quando o memorial de cálculo é desenvolvido em planilhas eletrônicas o ciclo de vida das questões tende a se reduzir ao tempo necessário para sua resolução (tema a ser abordado em trabalhos futuros).

Agradecimentos

Agradecemos ao apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES).

Referências

- Belhot, R. V. (2010). Taxonomia de Bloom : revisão teórica e apresentação das adequações do instrumento para definição de objetivos instrucionais. 421–431.
- Bloom, B. S., & Krathwohl, D. R. (1956). Taxonomy of educational objectives : the classification of educational goals. Handbook I, cognitive domain. Longmans.
- Canto, A. B. do, Ferreira, L., & Bercht, M. (2012). Objetos de aprendizagem no apoio à aprendizagem de engenharia: explorando a motivação extrínseca. *RENOTE- Revista Novas Tecnologias Na Educação*, 10(3). <http://www.seer.ufrgs.br/renote/article/view/36390/0>
- Canto Filho, A. B. do. (2022). Fábrica de questões: conceitos e métodos para o desenvolvimento de bancos de questões calculadas. <https://www.lume.ufrgs.br/handle/10183/243101>
- da Fonseca, M. A. R., Lima, D. da C. B. P., & de Oliveira Machado, E. C. (n.d.). REFERENCIAIS DE QUALIDADE PARA EaD: TRAJETÓRIA E INFLEXÕES NA EDUCAÇÃO SUPERIOR. Retrieved July 30, 2022, from <https://www.seminariosregionaisnape.net.br/numero7/Comunicacao/EIXO2PDF/23MariaAparecidaRodriguesDaFonsecaEDanielaDaCostaBrittoPereiralimaEElkaCandidaDeOliveiraMachadoE2.pdf>
- da Silva, I. V., & Afonso, A. F. (2021). Avaliação da aprendizagem em química: debates necessários no contexto de (pós) pandemia. *Research, Society and Development*, 10(9), e45310918111–e45310918111.
- de Souza, J. A., & Innocenti, M. S. (2022). Avaliação da aprendizagem escolar em matemática no contexto remoto: reflexões a partir de dizeres docentes. *Revista Meta: Avaliação*, 14(43), 292–315. <https://revistas.cesgranrio.org.br/index.php/metaavaliacao/article/view/3669/pdf>
- do Canto Filho, A. B. (n.d.). Desenvolvimento Sistemático de Questões Calculadas. <https://doi.org/10.13140/RG.2.2.32236.39044>
- Gagné, M., & Deci, E. L. (2005). Self-determination theory and work motivation. *Journal of Organizational Behavior*, 26(4), 331–362. <http://onlinelibrary.wiley.com/doi/10.1002/job.322/full>
- Herrera, M. M., & Angeletti, V. C. G. (2021). Evaluar en el contexto pandémico: hacia la evaluación con-formativa. *Revista Iberoamericana de Tecnología En Educación y Educación En Tecnología*, 28, e54–e54. <https://teyet-revista.info.unlp.edu.ar/TEyET/article/view/1549>
- Krathwohl, D. R. (2002). A revision of Bloom’s taxonomy: An overview. *Theory into Practice*, 41(4), 212–218.
- mac Gaul, M. I., Vargas, C. A., & del Olmo, P. (2021). Autogestión de la evaluación virtual en el campo de la Programación. *Revista Iberoamericana de Tecnología En Educación y Educación En Tecnología*, 28, e13–e13.
- Molon, J., Nicolao, M., & Franco, S. R. K. (2020). Ferramentas digitais para a avaliação do processo de aprendizagem: um mapeamento sistemático da literatura.

RENOTE, 18(2), 501–510. <https://www.seer.ufrgs.br/renote/article/download/110290/60045>

Moodle. (2022a). Moode - Calculated question type. https://docs.moodle.org/400/en/Calculated_question_type

Moodle. (2022b). Moodle Calculated Question Synchronization. https://docs.moodle.org/400/en/Calculated_question_type#Synchronization

Moodle Random Question. (2022). Moodle. https://docs.moodle.org/311/en/Random_question_type

Nigam, A., Pasricha, R., Singh, T., & Churi, P. (2021). A systematic review on AI-based proctoring systems: Past, present and future. *Education and Information Technologies*, 26(5), 6421–6445. <https://link.springer.com/article/10.1007/s10639-021-10597-x>

Pacheco, J. N., da Rosa, C. T. W., & Darroz, L. M. (2022). Motivação intrínseca, extrínseca e autoeficácia em relação à ciência e às aulas de ciências: estudo envolvendo os anos finais do Ensino Fundamental. *Research, Society and Development*, 11(5), e31111528380–e31111528380. <https://rsdjournal.org/index.php/rsd/article/download/28380/24598>

Pereira, A. S., Shitsuka, D. M., Parreira, F. J., & Shitsuka, R. (2018). Metodologia da pesquisa científica.

Presidência da República. (2017, May 25). DECRETO No 9.057, DE 25 DE MAIO DE 2017; diretrizes e bases da educação nacional.

Tuchtenhagen, P. ia, Tuchtenhagen, P. H., Heuert, S. K., dos Santos Dullius, A. I., Priesnitz, M. C., Ansuaj, A., Dullius, R. L., & Delboni, M. C. C. (2022). Abordagens de ensino e o processo avaliativo de alunos e professores. *Research, Society and Development*, 11(12), e347111234433–e347111234433.

Vansteenkiste, M., Lens, W., & Deci, E. L. (2006). Intrinsic Versus Extrinsic Goal Contents in Self-Determination Theory: Another Look at the Quality of Academic Motivation. *Educational Psychologist*, 41(1), 19–31. https://doi.org/10.1207/s15326985ep4101_4