

UNIVERSIDADE JOSÉ DO ROSÁRIO VELLANO - UNIFENAS
Merilza Rodrigues

COMPARAÇÃO ENTRE TESTES ESCRITOS E TESTES REALIZADOS NO
COMPUTADOR: ESTRESSE, SATISFAÇÃO E DESEMPENHO DOS ALUNOS DO
CURSO DE MEDICINA DA UNIFENAS – ALFENAS, CURSANDO GINECOLOGIA E
OBSTETRÍCIA

Belo Horizonte

2019

Merilza Rodrigues

COMPARAÇÃO ENTRE TESTES ESCRITOS E TESTES REALIZADOS NO
COMPUTADOR: ESTRESSE, SATISFAÇÃO E DESEMPENHO DOS ALUNOS DO
CURSO DE MEDICINA DA UNIFENAS – ALFENAS, CURSANDO GINECOLOGIA E
OBSTETRÍCIA

Dissertação apresentada ao curso de Mestrado
Profissional em Ensino em Saúde da Universidade José
do Rosário Vellano para a obtenção do título de Mestre
em Ensino em Saúde

Orientadora: Dra Maria Aparecida Turci

Belo Horizonte

2019

Ficha catalográfica elaborada pela Biblioteca Itapoã
Conforme os padrões do Código de Catalogação Anglo Americano (AACR2)

61-057

R696c

Rodrigues, Merilza.

Comparação entre testes escritos e testes realizados no computador: estresse, satisfação e desempenho dos alunos do curso de Medicina da Unifenas – Alfenas, cursando Ginecologia e Obstetria [manuscrito] / Merilza Rodrigues. -- Belo Horizonte, 2019.

59p. : il.

Dissertação (Mestrado) - Universidade José do Rosário Vellano, Programa de Pós-Graduação Stricto Sensu em Ensino em Saúde, 2019.

Orientador : Prof. Dra Maria Aparecida Turci.

1. Educação Médica. 2. Avaliação. 3. Computador. 4. Avaliação por Computador. 5. Avaliação aprimorada pelo computador. I. Turci, Maria Aparecida. II. Título.

Bibliotecária responsável: Kely A. Alves CRB6/2401



Presidente da Fundação Mantenedora - FETA

Larissa Araújo Velano Dozza

Reitora

Maria do Rosário Velano

Vice-Reitora

Viviane Araújo Velano Cassis

Pró-Reitor Acadêmico

Mário Sérgio Oliveira Swerts

Pró-Reitora Administrativo-Financeira

Larissa Araújo Velano Dozza

Pró-Reitora de Planejamento e Desenvolvimento

Viviane Araújo Velano Cassis

Diretor de Pesquisa e Pós-graduação

Mário Sérgio Oliveira Swerts

Vice-diretora de Pesquisa e Pós Graduação

Laura Helena Órfão

Coordenador do Curso de Mestrado Profissional em Ensino em Saúde

Antonio Carlos de Castro Toledo Jr.

Certificado de Aprovação

“COMPARAÇÃO ENTRE TESTES ESCRITOS E TESTES REALIZADOS NO COMPUTADOR:
ESTRESSE, SATISFAÇÃO E DESEMPENHO DOS ALUNOS DO CURSO DE MEDICINA DA
UNIFENAS – ALFENAS, CURSANDO GINECOLOGIA E OBSTETRÍCIA”

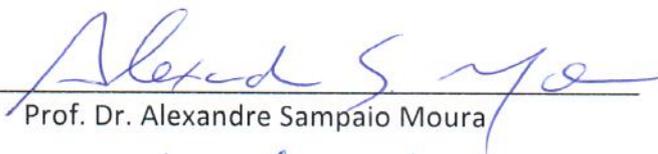
AUTOR: Merilza Rodrigues

ORIENTADOR: Profa. Dra. Maria Aparecida Turci

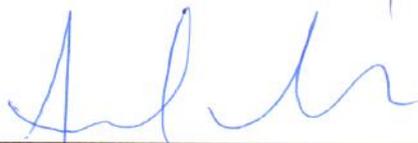
Aprovado como parte das exigências para obtenção do Título de **Mestre Profissional em Ensino em Saúde** pela Comissão Examinadora.



Profa. Dra. Maria Aparecida Turci



Prof. Dr. Alexandre Sampaio Moura



Prof. Dr. Gabriel Costa Osanan

Belo Horizonte, 28 de março de 2019.



Prof. Dr. Antonio Carlos de Castro Toledo Jr.
Coordenador do Mestrado Profissional
Em Ensino em Saúde
UNIFENAS

RESUMO

Introdução: As avaliações usando tecnologia estão presentes nas escolas em vários países, e seu uso se popularizou nas últimas duas décadas, sendo utilizadas no exame para licença médica nos Estados Unidos desde 1999. O potencial que as avaliações aprimoradas por tecnologia trazem para o mundo acadêmico foi o que motivou a realização deste estudo, e o conhecimento advindo de uma conferência específica sobre essas tecnologias trouxe o incentivo para persistir nessa linha de pesquisa. **Método:** Este estudo tem por objetivo comparar o estresse, a satisfação e o desempenho dos acadêmicos durante o primeiro semestre de 2018, ao realizarem provas escritas convencionais e provas no computador na disciplina de Ginecologia e Obstetrícia do curso de Medicina da Universidade José do Rosário Vellano, campus Alfenas. Foram avaliadas as turmas do décimo segundo período e sétimo período, totalizando uma amostra de 135 alunos. Tratou-se de um estudo experimental randomizado em que os alunos foram divididos em dois grupos para as avaliações e foram-lhes aplicados questionários pré-definidos e presenciais. Antes da prova, foram entregues aos alunos perguntas específicas para avaliar o estresse envolvido em sua realização. Ao término da prova, foram aplicadas as questões que avaliaram a satisfação com o tipo de prova realizada. Foram ainda registradas as notas obtidas nos dois tipos de avaliação. Cada aluno fez uma avaliação escrita convencional e uma avaliação via computador. **Resultados:** Os acadêmicos avaliados demonstraram maior estresse e menor satisfação com a prova no computador, comparando-a com a prova escrita convencional. O tipo de prova não esteve associado ao desempenho acadêmico. **Conclusão:** A falta de familiaridade e treinamento com a metodologia, antes da aplicação das avaliações, causou aumento no estresse nos acadêmicos envolvidos no presente estudo, o que interferiu na satisfação, ao fazer a prova usando tecnologia, embora não tenha havido diferença nos escores de desempenho. As informações coletadas serão utilizadas para reestruturação do formato dos testes oferecidos no departamento de Ginecologia e Obstetrícia da UNIFENAS, campus Alfenas, com o intuito de aprimorar a avaliação teórica que se realiza na Universidade.

Palavras Chave: Avaliação. Computador. Estudantes de Medicina. Avaliação por computador. Avaliação aprimorada pelo computador.

ABSTRACT

Introduction: Evaluations using technology are present in schools in several countries and have become popular in the last two decades and have been used in the United States medical license examination since 1999. The potential that improved technology assessments bring to the academic world motivated the carrying out of this study, and the knowledge from a specific conference on these technologies brought the incentive to persist in this research line. **Method:** This study aims to compare the stress, satisfaction and performance of academics during the first semester of 2018, when conducting conventional written tests and computer tests in the course of Gynecology and Obstetrics of the medical program of the University of José do Rosário Vellano, Alfenas campus, a total sample of 135 students. It was a randomized experimental study in which the students were divided for the evaluations and pre-defined and class-based questionnaires were applied. Prior to the test, specific questions were given to students to assess the stress involved in their achievement. At the end of the test, the questions were answered to evaluate the satisfaction with the type of the test they took. Each student did a conventional written assessment and a computer assessment. **Results:** The evaluated students showed greater stress and less satisfaction with the computer-based test compared to the conventional written test, although there was no statistical difference in performance when comparing both. **Conclusion:** Lack of familiarity and training with the methodology prior to test application resulted an increase in students' stress in the present study, which interfered in the satisfaction with the test using technology, even though there was no difference in the performance scores. The collected information will be used to restructure the format of the tests offered in the Department of Gynecology and Obstetrics of UNIFENAS, Campus Alfenas, in order to improve the theoretical evaluation performed at the University.

Key Words: Assessment. Computer. Medical Students. Computer Based Test. Technology Enhanced Evaluation. Computer assisted assessment.

LISTA DE ABREVIATURAS

HUAV	Hospital Universitário Alzira Vellano
CEP	Centro de Estudos e Pesquisas
SGA	Síndrome geral de adaptação
TCLE	Termo do consentimento livre e esclarecido
UNIFENAS	Universidade de Alfenas

LISTA DE TABELAS

Tabela 1	- Características sócio demográficas dos alunos, segundo Grupo alocado.....	27
Tabela 2	- Escores de Estresse e de Satisfação na 1ª fase do estudo, por Grupos (E/C e C/E) e períodos.....	29
Tabela 3	- Análise descritiva e comparativa entre os alunos que fizeram a prova no Computador e Escrita quanto às medidas dos escores de Estresse e de Satisfação na 1ª e 2ª fases do estudo.....	30
Tabela 4	- Estresse e de Satisfação e nota de desempenho considerando o tipo de prova respondida na 1ª fase.....	32
Tabela 5	- Análise de Correlação entre Escore de Estresse, Escore de Satisfação e Nota de desempenho avaliada nas duas fases do estudo, por tipo de prova.....	35
Tabela 6	- Análise de Correlação entre Escore de Estresse, Escore de Satisfação e Nota de desempenho avaliada independentemente da fase do estudo, por tipo de prova.....	36
Tabela 7	- Medida de confiabilidade da consistência interna das questões pertencentes aos escores de interesse – Prova no Computador.....	36
Tabela 8	- Medida de confiabilidade da consistência interna das questões pertencentes aos escores de interesse – Prova Escrita.....	38

LISTA DE ILUSTRAÇÕES

Figura 1	- Desenho esquemático dos agrupamentos realizados.....	21
Figura 2	- Parâmetros de interpretação.....	26
Gráfico 1	- Boxplot dos alunos que fizeram a prova no Computador e Escrita quanto às medidas dos escores de Estresse e de Satisfação – 1ª fase do estudo.....	31
Gráfico 2	- Boxplot dos alunos que fizeram a prova no Computador e Escrita quanto às medidas dos escores de Estresse e de Satisfação – 2ª fase do estudo.....	31
Gráfico 3	- Boxplot dos Escores de Estresse referente à comparação entre os 2 tipos de provas considerando o tipo de prova respondida na 1ª fase.....	33
Gráfico 4	- Boxplot dos Escores de Satisfação referente à comparação entre os 2 tipos de provas considerando o tipo de prova respondida na 1ª fase.....	34
Gráfico 5	- Boxplot das notas de desempenho dos alunos referente à comparação entre os 2 tipos de provas considerando o tipo de prova respondida na 1ª fase.....	34

SUMÁRIO

1	INTRODUÇÃO.....	9
2	JUSTIFICATIVA	17
3	OBJETIVOS	18
3.1	Objetivo geral.....	18
3.2	Objetivos específicos.....	18
4	MATERIAIS E MÉTODOS	19
4.1	Desenho do estudo.....	19
4.2	População.....	19
4.3	Critérios de inclusão	19
4.4	Critérios de exclusão.....	20
4.5	Amostragem e recrutamento	20
4.6	Elaboração das provas	20
4.7	Instrumentos utilizados	23
4.8	Variáveis	24
4.8.1	<i>Variáveis dependentes.....</i>	24
4.8.2	<i>Variáveis independentes</i>	25
4.9	Análise estatística.....	25
4.10	Aspectos éticos.....	26
5	RESULTADOS	27
6	DISCUSSÃO	39
7	CONCLUSÃO.....	44
	REFERÊNCIAS.....	45
	APÊNDICES	488

1 INTRODUÇÃO

As avaliações, de forma geral, e em especial de estudantes de Medicina, possuem vários propósitos, dentre eles, o de assegurar ao público que as competências necessárias foram aferidas, são adequadas e compatíveis ao profissional que a sociedade espera receber. A competência de um médico não inclui apenas o conhecimento específico, habilidades e raciocínio clínicos, mas também profissionalismo, estratégias de aprendizagem, utilização de recursos e capacidade de trabalho em equipe (APPEL et al., 2002). A esse respeito Epstein (2007) afirma ser a competência em Medicina o uso habitual e judicioso de comunicação, conhecimento, habilidades, raciocínio clínico, emoções, valores e reflexão na prática diária para o benefício dos indivíduos e comunidades que estão sendo atendidos. Posições como essas incitam ideias de como avaliar os graduandos para que sejam capazes e merecedores do que se costuma definir por competência, aumentando a responsabilidade pela formação e avaliação desse aluno.

Nesse contexto, a avaliação formativa e a capacitação dos professores, para oferecer *feedback* adequado, frequente e de qualidade, são fundamentais na formação dos futuros médicos (BORGES et al., 2014; CANOLE, 2005).

Na perspectiva da educação, avaliação é um termo que abrange qualquer atividade em que a evidência de aprendizagem é reconhecida de forma planejada e sistemática, sendo utilizada para emitir juízo sobre a aprendizagem. Os atributos da avaliação referem-se a sua função (PANUNCIO-PINTO; TRONCON, 2014). A finalidade da avaliação é, pois, auxiliar nas decisões sobre a melhor forma de dinamizar a aprendizagem, o que se costuma denominar avaliação formativa. Caso o objetivo consista em sintetizar a aprendizagem adquirida até o momento da avaliação, para certificar ou registrar o progresso, tem-se a avaliação somativa.

A escolha dos métodos a serem empregados na avaliação do estudante deve-se pautar pelo critério do melhor ajuste à natureza das habilidades e competências cujo domínio se quer conhecer. Não se pode deixar de considerar a importância da característica da avaliação, para se definir o diagnóstico prévio do conhecimento e melhor adequar suas necessidades ao contexto da disciplina. Essa tarefa é facilitada por alguns modelos conceituais, como a pirâmide de Miller, em que o aprendiz sabe, e sabe como, mostra e, enfim, faz; o que caracteriza uma avaliação formativa adequada. Ainda assim, ressalta-se que não existe uma maneira única de

avaliar que seja capaz de fornecer todos os dados e informações das habilidades e competências médicas (APPEL et al., 2002; CANOLE, 2005).

A avaliação, no caso específico da aprendizagem, subsidia um curso de ação para que se construa um resultado definido (FERREIRA, 2018). Para isso, é imperativo que exista uma ferramenta que possibilite ao docente organizar suas avaliações e gerenciar o desempenho e progresso dos seus educandos, um processo trabalhoso e que consome muitas horas, entretanto o uso da tecnologia é capaz de minimizar, em benefício de outras atividades.

O ato de avaliar traz uma multiplicidade de significados e interpretações: examinar, testar, distinguir, comparar e outros. É por essa variedade de abordagens que o processo de avaliar se torna tão difícil, ficando à mercê do julgamento de quem o pratica (PANUNCIO-PINTO; TRONCON, 2014). Nem sempre aqueles que se dedicam à docência da Medicina estão pedagogicamente preparados para o processo intrincado do ensino-aprendizagem. São habilitados como especialistas, atualizados em suas áreas, mas não são formados para serem professores e, menos ainda, avaliadores. Quando colocados nessa função, sentimentos contraditórios de competência e despreparo coexistem. Repetir os modelos que foram usados na sua formação se torna a estratégia mais segura, e só mesmo a busca por atualizações trará o conforto de ensinar e avaliar, fazendo com que se sintam mais aptos para desempenhar essa função. Por isso, o tema da avaliação assume papel de destaque nas discussões pedagógicas, em diversos cenários de ensino em que se buscam metodologias mais eficientes e melhor padronização das diferentes formas de avaliação (BORGES et al., 2014).

A tradição avaliativa do ensino brasileiro coloca ênfase no aspecto somativo, pois há necessidade de se produzir um resultado numérico, traduzindo nessa pontuação o critério de aptidão ou não do estudante, bem como registrar seu desempenho ao longo do curso e que, de certa forma, servirá de termômetro para guiá-lo na busca de oportunidades em sua formação (BORGES et al., 2014).

Na atualidade, os avaliadores têm a seu favor recursos tecnológicos que lhes garantem mais eficiência e dispõem de recursos que permitem fazer a transição dos modelos tradicionais para formas mais práticas, mais propícias a responder aos anseios dos estudantes que pertencem à geração tecnológica, por nascimento em um mundo conectado.

Historicamente, os computadores foram introduzidos no mundo acadêmico a partir da década de 1960 (WIJERANTHNE; RATHNAYAKE, 2013), mas seu uso se restringia aos grandes centros e às universidades dos países do primeiro mundo. O tamanho dos primeiros computadores, com pouca memória, e principalmente o alto custo das máquinas dificultavam o acesso ao grande público. Quando os computadores passaram a ser pessoais - chamados P.C. ou *personal computer* -, as máquinas mais acessíveis, logo se popularizaram, e os notebooks invadiram as universidades, facilitando o acesso à pesquisa.

No censo de 2010 (INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA, 2010), a presença do computador passou a fazer parte das questões formuladas pelo Instituto Brasileiro de Geografia e Estatística (IBGE), como aspecto de qualidade de vida, constando de seu instrumento investigativo a pergunta “se havia computadores em casa e se a família possuía acesso à internet”. Enquanto 90% da população brasileira possui geladeira e televisores e, embora a quantidade de computadores tenha mais que triplicado desde o censo de 2000, sua presença não chega a 40% dos lares brasileiros, e o acesso à internet é de 30,7%. Os microcomputadores estão presentes em 48% dos domicílios do Sudeste e em 19,4% nos do Norte. E em referência aos computadores com acesso à internet, no Sudeste, contabilizavam-se 39,6% e, no Norte, 15,4% (INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA, 2010). Comparando-se à Inglaterra, Hargreaves et al. (2004) observaram que, em 2002, 99% das crianças inglesas das escolas primárias usavam computador em casa, e 64% das casas tinham acesso à internet. Nessa época havia um computador para cada 9,7 crianças, uma realidade muito diferente da dos brasileiros, retratada na pesquisa do IBGE. No entanto, é possível inferir que o novo Censo irá indicar outra situação, devido às facilidades de acesso e aquisição de equipamentos, tanto por parte das famílias, como por parte das escolas privadas e públicas. Os indicadores ingleses demonstram por que os testes usando tecnologia fazem parte da rotina do país inglês há duas décadas.

Collins e Halverson (2010) afirmam que se está atravessando uma grande revolução na educação: a tecnologia digital, presente nos computadores e demais mídias, está modificando a visão de ensino e escolarização. A revolução industrial mudou o sistema de escolarização universal, já a tecnologia digital está no lado oposto desse modelo. Até as escolas para formação de condutores - as autoescolas - estão usando simuladores, e os videogames colocam os usuários em situações de desafios crescentes, como uma avaliação baseada no desempenho, o que pode ser levado para o ambiente educacional. Sem falar nos cursos online, que fazem questionar

como será a educação no futuro, sendo impossível persistir com a noção de aprendizado com a mesma fórmula para todos.

Em relação às pesquisas, o uso de tecnologia faz parte do cotidiano das escolas em todos os países, mas no que se refere ao uso dessa tecnologia como meio de avaliação, a história é mais recente. Como instrumentos de avaliação, os testes em computador foram introduzidos no exame para licença médica nos Estados Unidos a partir de 1999, conforme observam Cantillon, Irish e Sales (2004) e Wijerathne e Rathnayak (2013). Foram utilizados devido à preocupação com a segurança dos papéis impressos em relação aos dados coletados via computador, além do espaço físico e tempo requerido nesse processo. Vários autores, dentre os quais Clariana e Wallace (2002), confirmam a crescente utilização das avaliações realizadas em computadores, até mesmo para obter a licença para dirigir, exames para admissão em empresas do setor privado e certificações para grupos profissionais especializados.

Encontram-se na literatura atual, as siglas CAA, *computer assisted assessment*, ou ainda CBA, *computer based assessment*, para denominar as avaliações usando computadores. Outros termos que podem ser encontrados na literatura são *web based assessment*, *online assessment* e *computer aided assessment*, e o mais atual, TEA, *technology enhanced assessment*. Embora a terminologia e as interpretações individuais possam variar, as estratégias e práticas que possibilitam a eficácia do uso dos computadores são as mesmas (BULL; DANSON, 2004).

Computer assisted assessment é um termo amplo, que descreve as tecnologias envolvidas nos processos avaliativos usando os computadores. Incluem uma variedade de atividades para se acessar a informação. Sendo que CAA pode ser utilizada para avaliações formativas e ou somativas (BULL; DANSON, 2004).

Uma estratégia importante consiste em utilizar a metodologia oferecida pelo computador também para fazer um diagnóstico do conhecimento prévio que o aluno possui ao iniciar o curso, permitindo adequações do conteúdo e melhora do rendimento (CHALMERS; MCAUSLAND, 2002; CANOLE, 2005). Embora a avaliação diagnóstica não seja uma prerrogativa do computador, a análise dos resultados torna-se mais rápida com a análise estatística ao término da mesma.

Qualquer mudança exige adaptações e várias são as dificuldades para que se elaborem bancos de questões de qualidade e padronizados. As primeiras provas, constituídas de simples testes objetivos com alternativas, hoje podem ser ricas visualmente, porque além das imagens, pode-se utilizar vídeos curtos (CANTILLON; IRISH; SALES, 2004).

As vantagens iniciais com que os professores se deparam, ao usarem o computador, compreendem as pedagógicas e as administrativas. Do ponto de vista pedagógico, pode-se avaliar maior quantidade de tópicos com rapidez, possibilitando mudanças e avaliações do progresso do estudante. O tempo que se usa para correções e entrega de resultados deixa de existir, e o *feedback* imediato auxilia o aluno a melhorar sua performance (CHALMERS; MCAUSLAND, 2002). As vantagens administrativas incluem ganho de tempo em correções, diminuição do erro humano nas correções e as dificuldades para marcação da resposta em um gabarito. Se houver necessidade de se avaliar um grande grupo, economizam-se tempo e recursos. Por outro lado, levando-se em consideração grandes grupos, a economia em impressão é substancial, principalmente quando é necessário fazer um *upgrade* ou alterações em questões, fato comum na prática acadêmica. Além disso, avaliações estatísticas dos resultados podem ser realizadas em tempo real, caso seja necessário, até mesmo para avaliar a qualidade das questões, bem como o desempenho individual e do grupo.

Outra vantagem a ser apontada é a integração entre os setores acadêmicos e administrativos em relação aos resultados dos estudantes. Esse cenário de ganho bilateral apresenta grande potencial diagnóstico, melhorando de forma global as avaliações, sejam elas formativas ou somativas (CANTILLON; IRISH; SALES; 2004; BULL; DANSON, 2004; CHALMERS; MCAUSLAND, 2002).

Nas avaliações formativas, o aluno tem a opção de trabalhar sem supervisão e no seu próprio ritmo, podendo receber feedback do seu progresso e repetir as questões várias vezes, até a total compreensão do assunto. Já, nos testes somativos, são necessárias supervisão e limitação de tempo para execução e resultados (CHALMERS; MCAUSLAND, 2002).

A maior desvantagem do método de avaliação por computador é o tempo gasto na elaboração, correção e testagem das questões que irão compor o banco de questões. Canole e Warburton (2005) consideram erro pensar que as questões das avaliações escritas possam ser simplesmente migradas, para uma avaliação, usando tecnologia, é necessária uma adaptação, entre os

métodos, um processo contínuo, que requer desenvolvimento e treinamento dos professores nos aspectos pedagógicos e operacionais da tecnologia de CAA. É preciso confiar na segurança do sistema de infraestrutura e informática a sua disposição, principalmente quando se trata de avaliações somativas, em que os resultados precisam ser armazenados com segurança (BULL; DANSON, 2004; CHALMERS; MCAUSLAND, 2002).

Esses autores supracitados ressaltam ainda que, além do longo tempo necessário para a preparação acadêmica, um nível muito alto de cooperação e coordenação será necessário entre os departamentos. Dessa forma, várias fronteiras serão cruzadas dentro da instituição. Outro ponto a ser considerado são os estudantes com necessidades especiais, principalmente os deficientes visuais, que precisam de uma prova em *braille*, ou os disléxicos, que requerem um tempo maior (CHALMERS; MCAUSLAND, 2002).

Bull e Danson (2004) referem-se à necessidade de se atentar às adaptações a serem realizadas e os custos a serem considerados. Preocupações com a segurança são importantes, a randomização das questões é eficaz, sendo possível o uso de mais de uma resposta como correta ou incorreta, evitando comunicação entre os alunos. Muitas vezes, os laboratórios de informática das escolas mostram aparelhos com monitores altos, o que facilita ao aluno observar a tela ao lado. O acesso à internet é outro problema, porque um aluno poderia enviar a resposta para outro.

Cantillon, Irish e Sales (2004) afirmam que as questões foram ficando cada vez melhores com a evolução das tecnologias. Podendo-se lançar mão da reutilização das questões do banco, desde que esse seja rico e variado, e as perguntas demorem a se repetir. A variedade de testes que se pode usar no computador engloba questões dicotômicas, múltiplas escolhas, questões sequenciais, estudo de casos, questões com respostas curtas, questões usando imagem, tanto para pergunta como para resposta (CHALMERS; MCAUSLAND, 2002). Para se obter uma imagem de qualidade em testes no papel, o custo de uma impressão com boa definição limita o uso em grande escala.

Atualmente, é possível encontrar programas disponibilizados gratuitamente para desenvolver testes baseados em computador como o da University of Leicester's - CASTLE1. Essas facilidades não são apenas em relação aos testes. Os alunos podem ter acesso aos ambientes de aprendizagem virtual, cada vez mais comuns. Pacientes simulados como os desenvolvidos pela

Universidade de Miami Harvey Simulator estão mudando a forma de ensinar/aprender (CANTILLON; IRISH; SALES, 2004).

Vantagens práticas sobre os testes em papel são crescentes, mas a implementação dos testes *online* pode ser um desafio e deve levar algum tempo até sua aceitação geral. O conhecimento prévio de computação não parece interferir nos resultados para docentes ou discentes (LUECHT et al., 1998). Desde o início do CAA, o incremento da habilidade do uso de tecnologia tem facilitado a implantação desse método e pode-se lançar mão de um teste antes da avaliação verdadeira para dar tranquilidade aos alunos (BULL; DANSON, 2004).

Olgivie, Trusk e Blue (1999) sugerem que os exames administrados por computador são bem aceitos pelos acadêmicos de Medicina, pois acentuam o aprendizado, por meio da correção e *feedback* imediatos. Pode-se afirmar, pois, que avaliações baseadas em computador são tecnologia emergente, com potencial de promover avaliações de qualidade para médicos e outros profissionais da saúde.

A avaliação, independentemente da finalidade, é parte integrante da educação médica e serve a diversos fins, entre eles o feedback para alunos e professores, sobre o estado do conhecimento dos educandos. Além disso, é um meio para os alunos rastream sua aptidão e o resultado dos seus esforços na busca do conhecimento (CHALMERS; MCAUSLAND, 2002).

Ao usar a tecnologia para aplicação dos exames no computador é fundamental que a preocupação seja avaliar o estudante quanto ao conhecimento adquirido e não seu nível de conforto e confiança no uso da tecnologia.

Avaliações são consideradas partes estressantes e geradoras de alto grau de ansiedade, dentro da educação médica (YATES, 2012). A ansiedade tem um importante papel no processo ensino/aprendizagem. Se os professores puderem reconhecer em que situações os testes geram maior grau de ansiedade, esses testes poderão ser omitidos ou descartados. Dessa forma, os resultados das avaliações provavelmente irão aferir o conhecimento dos estudantes de forma mais cuidadosa (AZIMI, 2016) e irão possibilitar avaliar também a ansiedade relatada pelos acadêmicos na realização das provas.

O termo estresse foi utilizado pela primeira vez, com o contexto atual, por Hans Selye, em 1936, para designar o que chamou de “síndrome geral da adaptação” (SGA). Para ele, a SGA se desenvolvia em três fases: aguda, o corpo produz mais força, energia e motivação; fase de resistência, em que o corpo tenta se reestabelecer; fase de exaustão, o corpo volta aos sintomas da primeira fase, com mais intensidade (LIMA et al., 2016). Lipp (2000), em sua visão cognitivo-comportamental, distingue estresse saudável (eustresse), como um mecanismo de luta e fuga, que aumenta a produção e criatividade do indivíduo, e estresse patológico (distresse), que leva a respostas inadequadas, enfermidades físicas e psicológicas (LIMA et al., 2016).

O estresse nos estudantes de Medicina, originado por avaliações do conhecimento, é especialmente preocupante, pois, ao afetar funções fisiológicas, psicológicas e cognitivas, prejudica a qualidade de vida, influi no aprendizado e no cuidado ao paciente. (LIMA et al., 2016). Sim, Holifield e Brown (2004) encontraram em sua investigação, sobre implementação da avaliação por computador, o aprendizado superficial e o aumento da ansiedade, com a diminuição das notas, ao usar essa metodologia.

2 JUSTIFICATIVA

Existem lacunas em relação ao conhecimento do estresse, satisfação e desempenho dos alunos frente aos testes em computador, pois poucos são os relatos encontrados na literatura. Procuramos com esse estudo avaliar, junto aos acadêmicos de Medicina da Universidade José do Rosário Vellano, o impacto que o uso de tecnologia irá gerar nas provas realizadas por eles.

3 OBJETIVOS

3.1 Objetivo geral

Avaliar o estresse, desempenho e o nível de satisfação dos acadêmicos de Medicina do sétimo e décimo segundo períodos, ao serem submetidos a testes escritos convencionais e a testes em computador.

3.2 Objetivos específicos

- Caracterizar os participantes quanto às condições sócio-demográficas;
- Comparar o estresse vivenciado por acadêmicos de Medicina do sétimo e décimo segundo períodos cursando Ginecologia e Obstetrícia ao realizarem um teste escrito convencional e um teste no computador;
- Verificar o desempenho entre o teste escrito convencional e o realizado no computador por acadêmicos de Medicina;
- Comparar a satisfação obtida por acadêmicos de Medicina do sétimo e décimo segundo períodos, cursando a disciplina de Ginecologia e Obstetrícia, entre o teste escrito convencional e ao realizado no computador;
- Comparar as possíveis diferenças nos escores de estresse, desempenho e satisfação entre acadêmicos de Medicina de acordo com o tipo de prova.

4 MATERIAIS E MÉTODOS

4.1 Desenho do estudo

Trata-se de um estudo experimental randomizado em duas fases, que analisou grupos de alunos submetidos a testes escritos convencionais e testes em computador, avaliando seu estresse, desempenho e nível de satisfação.

4.2 População

Os participantes do estudo foram os alunos do curso de Medicina da Universidade José do Rosário Vellano (UNIFENAS) do Campus de Alfenas, matriculados no décimo segundo e sétimo períodos, somando aproximadamente 161 alunos, que cursaram a disciplina de Ginecologia e Obstetrícia no primeiro semestre de 2018.

Os acadêmicos têm em sua matriz curricular as disciplinas Ginecologia e Obstetrícia no sétimo período, sendo que o módulo se denomina Saúde da Mulher. Ginecologia e Obstetrícia 1 no décimo período, sendo que esse módulo do internato do quinto ano é feito em outro município, por isso não foram convidados a participarem do trabalho. No décimo segundo período, o módulo denomina-se Ginecologia e Obstetrícia II. As turmas subdivididas em três subgrupos cursaram o módulo de Ginecologia e Obstetrícia a cada sete semanas, período em que foram submetidos a duas avaliações somativas, uma escrita e outra no computador, com valor de dez pontos, distribuídos entre as duas avaliações.

4.3 Critérios de inclusão

Foram incluídos todos os acadêmicos do sétimo e décimo segundo períodos no primeiro semestre de 2018, do curso de Medicina da Universidade José do Rosário Vellano (UNIFENAS), do campus de Alfenas, que, após serem orientados sobre o estudo, assinaram o termo de consentimento livre e esclarecido.

4.4 Critérios de exclusão

Excluíram-se do estudo aqueles alunos que não quiseram participar do estudo e não assinaram o termo de consentimento, bem como os alunos que repetiram o módulo. Os alunos do terceiro módulo do décimo segundo período foram excluídos. Essa exclusão se deu, pois, pelo fato de esse módulo ter apenas cinco semanas de duração, e ser realizada apenas uma avaliação. Foram excluídos no total 26 acadêmicos.

4.5 Amostragem e recrutamento

Utilizou-se amostra por conveniência. Todos os alunos que cursaram a disciplina Ginecologia e Obstetrícia, no primeiro semestre de 2018, portanto cursando o sétimo e décimo segundo períodos de Medicina, foram convidados a participar do estudo, perfazendo um total de 161 alunos.

O recrutamento foi feito no primeiro dia do módulo de Ginecologia e Obstetrícia, no momento da apresentação do curso, quando o objetivo e a metodologia do estudo foram apresentados aos alunos.

O termo do consentimento livre e esclarecido (TCLE) (APÊNDICE C) foi entregue aos interessados em participar do estudo e/ou enviado por endereço eletrônico, através da secretaria acadêmica da Universidade. Aqueles que, após lerem o TCLE, concordaram em participar do estudo e assinaram o documento fizeram parte do estudo.

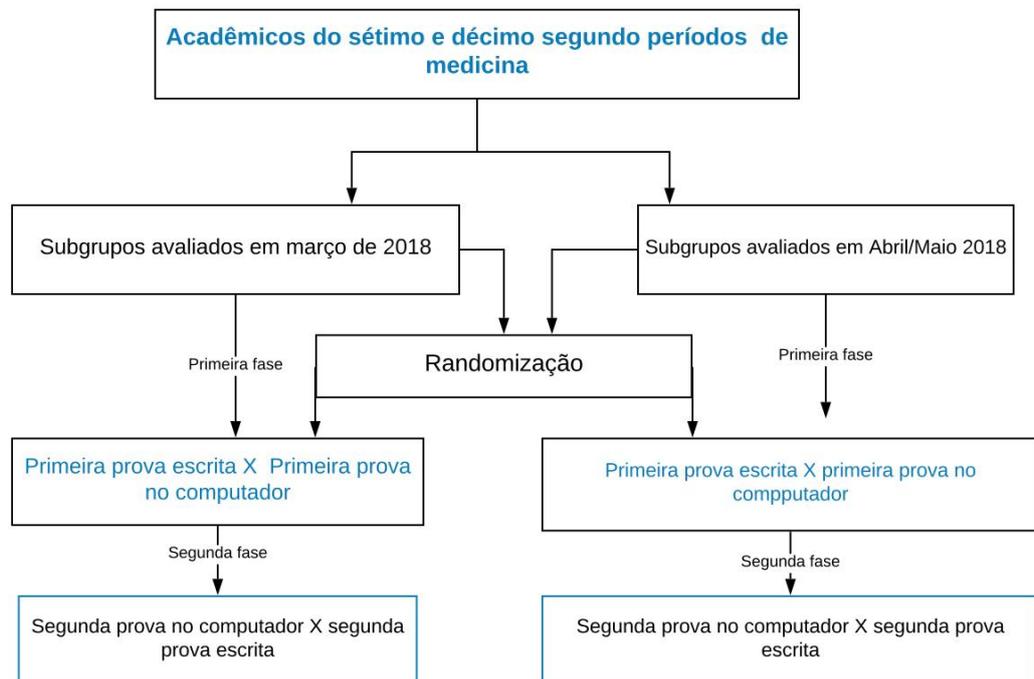
4.6 Elaboração das provas

O curso de Medicina da UNIFENAS, no campus de Alfenas, segue o sistema modular após o nível básico. Sendo assim, ao iniciarem o sétimo período, os alunos são divididos em três subgrupos e cursam, em sistema de rodízio dentro do semestre, os módulos “saúde da mulher”, “clínica médica” e “clínica cirúrgica”. Ao chegarem ao décimo período, cursam a disciplina Ginecologia e Obstetrícia 1, em cenários de prática fora do campus e por esse motivo não foram incluídos no estudo. Já no décimo segundo período cursam Ginecologia e Obstetrícia 2, Programa Saúde da Família e Urgência e Emergência. Cada módulo tem duração de sete semanas e, durante esse período, os alunos são avaliados em habilidades práticas (10 pontos),

conceito (satisfatório ou insatisfatório), com realização de duas avaliações teóricas (5 pontos cada uma), contendo 30 questões. Até então as provas sempre foram realizadas em papel. A somatória das duas avaliações cognitivas confere dez pontos. Essas duas provas teóricas foram avaliadas no presente estudo.

Os alunos dos sétimo e décimo segundo períodos foram distribuídos, portanto, em dois grupos: “prova escrita convencional/prova computadorizada” (Grupo E/C) e “prova computadorizada/prova escrita convencional” (Grupo C/E). Dessa forma, metade da turma do sétimo e metade da turma do décimo segundo períodos realizou a 1ª prova da disciplina como habitualmente se faz no curso - em papel - e a 2ª prova foi realizada no formato informatizado (Grupo E/C). Já o outro grupo de alunos fez o inverso, a primeira prova no computador e a segunda, em papel (Grupo C/E), conforme ilustrado da FIG. 1.

Figura 1 – Desenho esquemático dos agrupamentos realizados



Fonte: Dados referentes ao estudo

A randomização dos alunos nos dois grupos foi realizada de forma computadorizada, utilizando-se o programa Microsoft Excel para o Windows 2007 (Microsoft Corporation, Redmond, WA, EUA) e sua ordem era desconhecida dos investigadores. Após ser gerada por

um programa para computador, a lista de randomização foi mantida em sigilo. Para isso, foram utilizados códigos fornecidos por computador (*web-based randomization*).

As provas escritas convencionais foram realizadas em sala de aula teórica e as provas informatizadas, no laboratório de informática, ambas localizadas no prédio do centro de estudos e pesquisas, do Curso de Medicina da Universidade. A sala de aula teórica utilizada neste estudo localiza-se no mesmo andar do laboratório de informática, e as provas sempre foram simultâneas. O laboratório de informática do Centro de Estudos e Pesquisas (CEP) faz parte do complexo do Hospital Universitário Alzira Velano (HUAV). A sala possui 32 computadores alinhados de frente com 16 máquinas de cada lado. São do tipo Cpu – Laboratorio Pc 32 - Dual Core 2.70 Ghz - 2 Gb Ram - 160 GbDell Optiplex 360 Desktop U0254 S.O.: Windows 7. As provas foram realizadas simultaneamente e tiveram o mesmo conteúdo, com questões idênticas, e com o mesmo tempo de duração (uma hora). Foi utilizada a plataforma Moodle, disponível ao corpo docente da Universidade, primeiro para a execução do banco de questões e posteriormente para a realização das provas informatizadas.

Os alunos foram informados da sua alocação nos momentos que antecederam a primeira prova da disciplina. Somente os pesquisadores tiveram conhecimento da ordem do tipo de prova para cada grupo.

Antes de cada avaliação, todos os alunos responderam ao questionário específico de estresse em relação à prova, no computador ou escrita, e ao final das provas o questionário de satisfação (APÊNDICES A e B), relatando sua experiência com as duas avaliações. Ao final do módulo, compararam-se as notas entre as duas avaliações, para definição do desempenho dos acadêmicos.

Cada professor do Departamento de Ginecologia e Obstetrícia da UNIFENAS foi responsável por elaborar entre cinco e dez questões por assunto lecionado. Os professores foram incentivados a colocar imagens, caso possível, com o intuito de melhorar o *layout* das provas. Estimulou-se a clareza do enunciado, bem como a preferência por casos clínicos curtos. As questões foram padronizadas, contendo quatro alternativas com respostas corretas ou incorretas, dependendo do contexto. As mesmas questões foram utilizadas para o teste escrito e do computador. Foi criado, dessa forma, um banco de questões para que se pudesse lançar

mão da quantidade necessária para elaboração das avaliações, cuidando-se para que não houvesse repetição das questões. Antes desse estudo não havia um banco de questões.

Definiu-se para essa pesquisa, formatar as provas realizadas, seja no papel ou no computador, com 30 questões, que deveriam ser respondidas em uma hora. Levando-se em consideração que os alunos que fizeram a prova escrita, ao concluí-la, precisavam transcrever as respostas para o gabarito, um processo dispensável no computador, mas que leva tempo adicional.

4.7 Instrumentos utilizados

Aplicou-se questionário semiestruturado, em versões em papel e eletrônica, contendo 28 questões e dividido em três partes e com as seguintes informações (APÊNDICES A e B)

1) Dados sócio-demográficos e comportamentais: idade, sexo, estado civil, etnia, origem do ensino médio (escola pública ou privada), ano de ingresso no curso, realização ou não de trabalho remunerado (número de horas);

2) Estresse relacionado às provas, (adaptado do questionário de *stress* de Lipp): se sente frio na barriga antes das provas; coração bate mais rápido antes das provas; sente-se depressivo antes das provas, sente-se angustiado antes das provas; sente-se pressionado antes das provas; a boca seca antes das provas; precisa ir ao banheiro mais vezes antes das provas; se sente estressado antes das provas; sente-se mais estressado com a prova escrita/computador; se questiona sobre a suficiência do desempenho na realização da prova; fica preocupado com a possibilidade de reprovação na realização da prova; preocupa-se com a quantidade de conteúdo cobrado na realização da prova; confia no próprio desempenho na realização da prova; fica satisfeito consigo na realização da prova; sente-se mais confiante na prova escrita/computador.

3) Satisfação com as provas, (perguntas formuladas pela autora): adequação do formato da prova escrita/computador; atratividade do *layout* da prova escrita/computador, facilitando o entendimento e interpretação das questões; conforto em relação à prova escrita/computador e relação com a piora do desempenho; suficiência do tempo para a conclusão da prova; satisfação com a realização da prova escrita/computador.

Em cada fase do estudo os alunos responderam ao questionário de stress antes da realização de cada prova e o de satisfação ao seu término. Sendo que cada aluno respondeu ao questionário duas vezes.

Todas as avaliações, prova escrita e prova no computador, foram idênticas e tiveram o mesmo valor.

No caso da prova no computador, ao seu término, os académicos obtiveram o resultado imediato e puderam conferir o feedback disponível na correção de cada questão. A plataforma utilizada permite e facilita esse procedimento, assim como o envio à secretaria dos resultados com a estatística de acerto e erro de cada questão.

Aos alunos da prova escrita foi entregue o gabarito com a resposta das questões para sua conferência. As notas oficiais após correção foram disponibilizadas aos alunos, e os resultados entregues na secretaria para registro. O *feedback* foi dado em sala de aula, após a correção.

4.8 Variáveis

4.8.1 Variáveis dependentes

O estresse foi avaliado por questionário específico, baseado no Inventário de Sintomas de Estresse para Adultos de Lipp (2000) com 15 perguntas. Esse questionário de estresse foi aplicado antes da entrega das provas escritas ou no computador, e, só depois de respondido, ocorreu o início das provas.

A satisfação do aluno com o tipo de prova foi avaliada por meio de questionário com cinco perguntas, listadas na seção anterior. As respostas foram categorizadas, utilizando-se Escala Likert com 7 afirmativas, sendo 1, discordo totalmente até 7, concordo totalmente.

O desempenho na prova foi a última variável analisada e foi expressa pela porcentagem de acertos na avaliação de valor 5.0. Avaliou-se o próprio aluno em relação às duas provas, e os grupos entre si.

4.8.2 Variáveis independentes

As variáveis explicativas foram as seguintes: 1- o tipo de prova; 2- idade (em anos); 3- sexo (feminino e masculino); 4- estado civil (casado/ união estável, solteiro, separado); 5- Cor da pele (autorreferida e classificada em branca, negra, parda, amarela, indígena ou outra (BRASIL, 2000); 6- origem do ensino médio (escola pública ou privada); 7- ano de ingresso no curso; 8- realiza trabalho remunerado.

4.9 Análise estatística

Na análise descritiva, a amostra foi caracterizada a partir do cálculo da média e intervalo de 95% de confiança para a variável numérica e de percentuais, como medidas para descrever os resultados das variáveis categóricas.

O indicador de estresse foi calculado a partir da média das respostas às perguntas ($Q1 + Q2 + Q3 + Q4 + Q5 + Q6 + Q7 + Q8 + Q9 + Q10 + Q11 + Q12 + Q13_{inv} + Q14_{inv} + Q15_{inv} / 15$), assim como o de satisfação ($Q1 + Q2 + Q3_{inv} + Q4 + Q5 / 5$). As perguntas com a sigla INV tiveram o sentido da escala invertido para avaliações positivas e negativas.

Com o objetivo de comparar independentemente os dois grupos quanto às médias, utilizou-se o teste *t* de *Student* para amostras independentes. O teste de Levene foi utilizado com o objetivo de averiguar a homogeneidade das variâncias de cada item (variável) estudado, por grupo. Neste estudo, decidiu-se por assumir a heterogeneidade das variâncias; com isso, optou-se por utilizar os valores do teste *t* de *Student*, assumindo-se a não igualdade de variâncias, o que contribui para obterem-se resultados mais robustos. A diferença entre os grupos em relação às variáveis categóricas foi realizada utilizando-se o teste Qui-quadrado de Pearson. No caso de ocorrer associação significativa entre duas variáveis de interesse, em tabelas 2x2, avaliou-se também a estatística *Odds Ratio* (Razão das Chances \rightarrow O. R.), medida que quantifica o grau de associação entre duas variáveis categóricas. O teste exato de Fisher foi usado nos casos em que um número de casos inferiores a cinco ocorreram em tabelas 2x2, medida que quantifica o grau de associação entre duas variáveis categóricas. Na comparação entre medidas realizadas em um mesmo grupo, avaliadas em dois momentos distintos, o teste estatístico *t* de *Student* para amostras pareadas / dependentes foi utilizado.

A relação entre as variáveis de interesse foi feita, empregando-se a Correlação de Pearson, utilizando-se os seguintes parâmetros de interpretação (FIG. 2).

Figura 2 - Parâmetros de interpretação

Pontos de corte	“Interpretação”
$ r < 0,40$	Correlação Fraca
$0,40 \leq r \leq 0,70$	Correlação de Moderada a Forte
$ r > 0,70$	Correlação Forte

Fonte: Dados referentes ao estudo

A inconstância das variáveis entre grupos foi verificada utilizando-se o gráfico *Boxplot*.

A confiabilidade (Consistência interna) dos instrumentos utilizados foi analisada utilizando-se o coeficiente Alfa de Cronbach, que apresenta a média de todos os coeficientes de correlação possíveis, resultantes das diferentes divisões em metades, de acordo com as questões/afirmativas que compõem cada uma dos escores pesquisados, bem como possibilitam a avaliação do impacto da retirada de cada questão/afirmativa no conjunto de questões pertencente a um determinado escore de interesse (Domínio/Dimensão latente). Além disso, é avaliada a correlação item-total, medida que avalia a correlação do item (questão/afirmativa) com o escore a que pertence.

Todos os resultados foram considerados significativos para uma probabilidade de significância inferior a 5% ($p < 0,05$), tendo, portanto, pelo menos 95% de confiança nas conclusões apresentadas.

4.10 Aspectos éticos

Os acadêmicos foram convidados a participar do estudo por livre e espontânea vontade. O esclarecimento da pesquisa proposta foi explicitado pelo Termo de Consentimento Livre e Esclarecido (TCLE, Apêndice C). O estudo foi submetido e aprovado pelo comitê de ética em pesquisa (APÊNDICE E) da Universidade José do Rosário Vellano, e, somente após seu aval, com número do parecer: 2.535778, a pesquisa teve início.

5 RESULTADOS

Foram excluídos 26 alunos que não preencheram os critérios do trabalho, sendo incluídos nos resultados apresentados 135 acadêmicos: 84 alunos do sétimo período e 51 do décimo segundo período.

Os grupos foram homogeneamente distribuídos, com 66 alunos no Grupo E/C (41 do sétimo e 25 do décimo segundo) e 69 no Grupo C/E (43 alunos do sétimo e 26 do décimo segundo período). A amostra foi composta em sua maioria por mulheres, com idade média de 25 anos, pessoas solteiras e brancas, com origem no ensino privado e sem atividade remunerada. Os grupos foram comparados e verificou-se que eles não se diferiam em relação a essas características (TAB. 1).

Tabela 1 – Características sociodemográficas dos alunos, segundo grupo alocado
(Continua)

Variáveis	Grupo Comp/Escrita		Grupo Escrita/Comp.		Total	
	n	%	n	%	n	%
Período do curso						
<i>7º período</i>	43	62,3	41	62,1	84	62,2
<i>12 período</i>	26	37,3	25	37,9	51	37,8
	p = 0,981					
Sexo						
<i>Masculino</i>	28	40,6	20	30,3	48	35,6
<i>Feminino</i>	41	59,4	46	69,7	87	64,4
	p = 0,212					
Idade (anos)						
<i>Média (I.C. 95%)</i>	25,6 (24,6; 26,6)		24,9 ± 3,4 (24,0; 25,7)		25,3 ± 3,8 (24,6; 25,9)	
	p* = 0,252					
Estado civil						
<i>Solteiro(a)</i>	61	88,4	62	93,9	123	91,1
<i>Casado(a) / União estável</i>	8	11,6	4	6,1	12	8,9
	p = 0,259					

Tabela 1 – Características sociodemográficas dos alunos, segundo grupo alocado (conclusão)

Variáveis	Grupo Comp/Escrita		Grupo Escrita/Comp.		Total	
	n	%	n	%	n	%
Raça/Cor						
<i>Branca</i>	54	78,3	55	83,3	109	80,8
<i>Parda</i>	12	17,4	10	15,2	22	16,3
<i>Negra</i>	1	1,4	0	0,0	1	0,7
<i>Amarela</i>	2	2,9	1	1,5	3	2,2
p** = 0,894						
Origem do ensino médio						
<i>Público</i>	9	13,0	13	19,7	22	16,3
<i>Privado</i>	55	79,7	51	77,3	106	78,5
<i>Público e Privado</i>	5	7,3	2	3,0	7	5,2
p** = 0,343						
Trabalho remunerado						
<i>Sim</i>	3	4,3	0	0,0	3	2,2
<i>Não</i>	66	95,7	66	100,0	132	97,8
p** = 0,245						

Base de Dados: 135 alunos (*Comp./Escrita* → 69 alunos e *Escrita/Comp.* → 66 alunos) – 1ª prova

Nota: I.C. 95% → Intervalo de confiança de 95% da média.

p → Probabilidade de significância do teste *qui-quadrado de Pearson*

p* → Probabilidade de significância do teste *t de student para amostras independentes*

p** → Probabilidade de significância do teste *Exato de Fisher*

Fonte: Dados referentes ao estudo

Na sequência, foram verificados os escores de estresse e satisfação por grupos e períodos. A TAB. 2 apresenta as médias, valores mínimos, máximos e desvio padrão de cada subgrupo. Não existem diferenças entre os alunos do 7º e 12º períodos quanto aos escores de estresse e de satisfação, independentemente do tipo de prova a que o aluno tenha respondido (Computador ou Escrita), na 1ª prova aplicada aos alunos. Por esse motivo, as comparações subsequentes quanto aos escores de estresse e satisfação agruparão os períodos.

Tabela 2 - Escores de estresse e satisfação na 1ª fase do estudo, por grupos (E/C e C/E) e períodos.

Prova	Período	n	Mínimo	Máximo	Média	d.p.	p*
Grupo Computador/Escrita							
<i>Escore: Estresse</i>							
	7º	43	1,53	6,40	4,41	1,21	0,783
	12º	26	1,67	6,53	4,50	1,37	7º = 12º
<i>Escore: Satisfação</i>							
	7º	43	1,80	6,80	4,18	1,53	0,240
	12º	26	1,60	7,00	4,59	1,34	7º = 12º
Grupo Escrita/Computador							
<i>Escore: Estresse</i>							
	7º	41	2,13	6,73	4,42	1,31	0,354
	12º	25	1,47	6,13	4,15	1,03	7º = 12º
<i>Escore: Satisfação</i>							
	7º	41	1,00	7,00	5,17	1,36	0,465
	12º	25	2,80	7,00	5,41	1,26	7º = 12º

*O valor de p refere-se ao teste *t de student* para amostras independentes.

Fonte: Dados referentes ao estudo

Posteriormente foram comparados os escores de estresse e satisfação medidos na primeira e na segunda provas (TAB. 3). Quanto ao escore de estresse, não se verifica diferença entre os grupos que realizaram a prova no Computador e a prova Escrita, tendo sido essa prova realizada na 1ª ou 2ª fases do estudo. Já no que diz respeito ao escore de satisfação, os resultados indicam que os alunos que realizaram provas escritas apresentaram maior satisfação, independentemente de terem-se submetido a elas, na primeira ou na segunda provas (GRÁF. 1 e 2).

Tabela 3 - Análise descritiva e comparativa entre os alunos que fizeram a prova no Computador e Escrita quanto às medidas dos escores de Estresse e de Satisfação nas 1ª e 2ª fases do estudo.

	Prova	n	Mínimo	Máximo	Média	d.p.	p*
1ª fase do estudo							
<i>Escore: Estresse</i>							
	<i>Computador</i>	69	1,53	6,53	4,45	1,26	0,554
	<i>Escrita</i>	66	1,47	6,73	4,32	1,21	C = E
<i>Escore: Satisfação</i>							
	<i>Computador</i>	69	1,60	7,00	4,33	1,46	< 0,001
	<i>Escrita</i>	66	1,00	7,00	5,26	1,32	C < E
2ª fase do estudo							
<i>Escore: Estresse</i>							
	<i>Computador</i>	66	1,53	6,47	4,29	1,23	0,222
	<i>Escrita</i>	69	1,40	6,40	4,04	1,15	C = E
<i>Escore: Satisfação</i>							
	<i>Computador</i>	66	1,20	7,00	3,97	1,36	< 0,001
	<i>Escrita</i>	69	2,00	7,00	5,01	1,39	C < E

d.p. → Desvio-padrão.

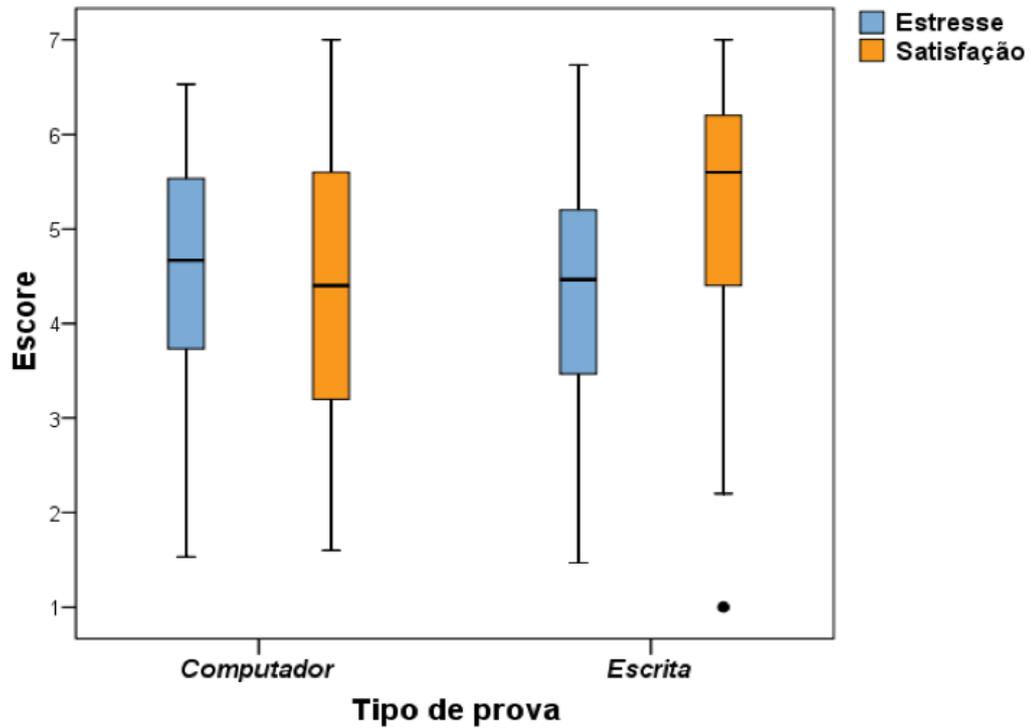
*O valor de p refere-se ao teste *t de student* para amostras independentes.

Legenda: C → Prova no Computador

E → Prova Escrita

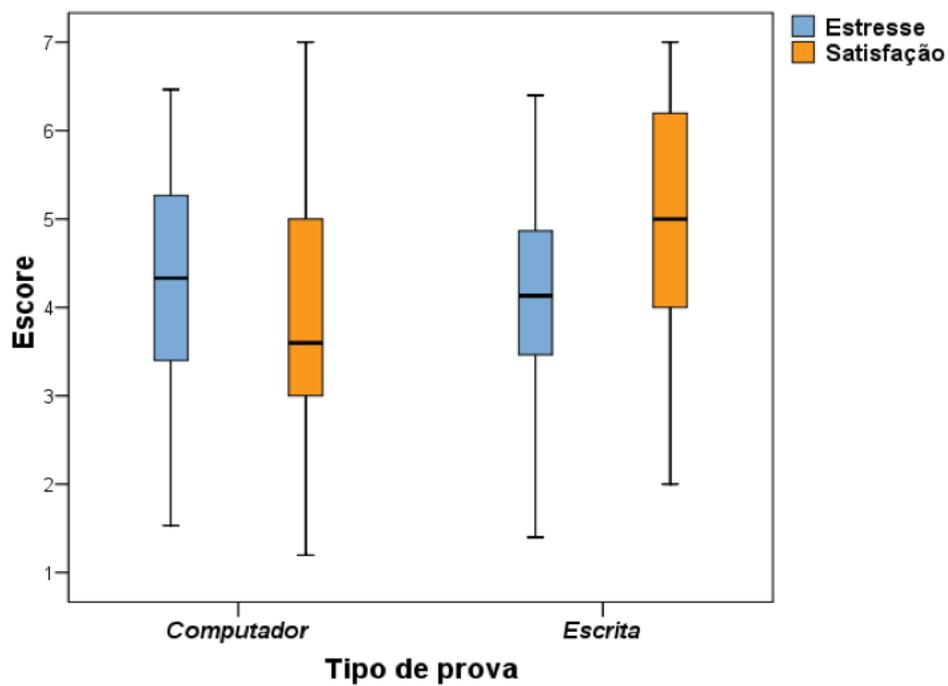
Fonte: Dados referentes ao estudo

Gráfico 1 - *Boxplot* dos escores de Estresse e de Satisfação referentes aos alunos que fizeram a prova no Computador e Escrita – 1ª fase do estudo.



Fonte: Dados referentes ao estudo

Gráfico 2 - *Boxplot* dos escores de Estresse e de Satisfação referentes aos alunos que fizeram a prova no Computador e Escrita – 2ª fase do estudo.



Fonte: Dados referentes ao estudo

Na comparação entre os dois tipos de provas aplicadas na 1ª e 2ª fases do estudo respondido pelo mesmo aluno (em ordem diferente de acordo com o tipo de prova respondido na 1ª fase do estudo), o escore de estresse dos respondentes da 1ª prova via computador e da 2ª prova, de forma escrita, foi maior na primeira prova. Já no grupo que respondeu à 1ª prova via computador e à 2ª prova de forma escrita, os escores de estresse foram semelhantes. No que se refere ao escore de satisfação, os resultados apontam que nos dois grupos (alunos que responderam a 1ª prova via computador e a 2ª prova de forma escrita e alunos que responderam a 1ª prova de forma escrita e a 2ª prova via computador), o escore de satisfação com a prova respondida via computador foi significativamente menor do que quando respondida via convencional (TAB. 4 e GRÁF. 3 e 4). A TAB. 4 contém, ainda, a comparação entre as notas de desempenho dos alunos, indicando que o desempenho dos alunos foi semelhante, independentemente do tipo de prova (GRÁF. 5).

Tabela 4 – Escores de estresse, satisfação e nota de desempenho considerando o tipo de prova respondida na 1ª fase

							(Continua)
Score	Prova	n	Mínimo	Máximo	Média	d.p.	p*
<u>Score: Estresse</u>							
1ª prova:							
Computador							
	<i>Computador</i>	69	1,53	6,53	4,45	1,26	< 0,001
	<i>Escrita</i>	69	1,40	6,40	4,04	1,15	C > E
1ª prova: Escrita							
	<i>Computador</i>	66	1,53	6,47	4,29	1,23	0,704
	<i>Escrita</i>	66	1,47	6,73	4,32	1,21	C = E
<u>Score: Satisfação</u>							
1ª prova:							
Computador							
	<i>Computador</i>	69	1,60	7,00	4,33	1,46	0,025
	<i>Escrita</i>	69	2,00	7,00	5,01	1,39	C < E

Tabela 4 – Escores de estresse, satisfação e nota de desempenho considerando o tipo de prova respondida na 1ª fase

Score	Prova	n	Mínimo	Máximo	Média	d.p.	(Conclusão) p*
1ª prova: Escrita							
	<i>Computador</i>	66	1,20	7,00	3,97	1,36	< 0,001
	<i>Escrita</i>	66	1,00	7,00	5,26	1,32	C < E
Nota de desempenho (%)							
1ª prova: Computador							
	<i>Computador</i>	69	43,00	93,25	71,10	10,64	0,888
	<i>Escrita</i>	69	40,00	90,00	70,90	10,24	C = E
1ª prova: Escrita							
	<i>Computador</i>	66	45,67	93,33	70,74	10,00	0,204
	<i>Escrita</i>	66	42,75	100,00	73,07	11,86	C = E

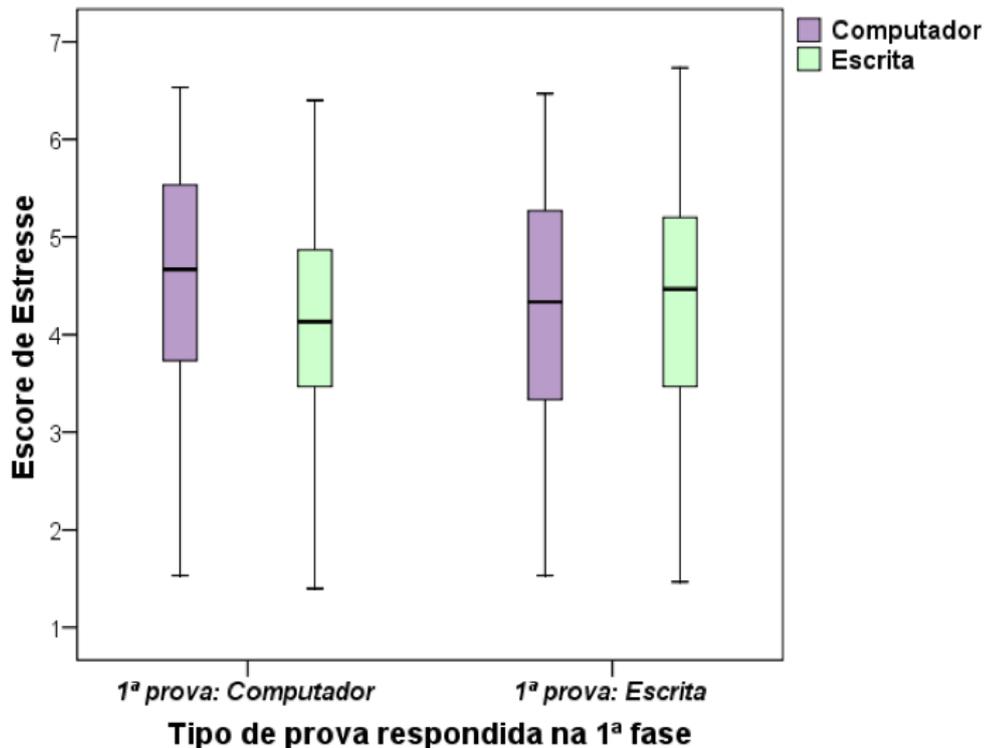
d.p. → Desvio-padrão

*O valor de p refere-se ao teste *t de student* para amostras pareadas.

legenda: C → Prova no Computador E → Prova Escrita

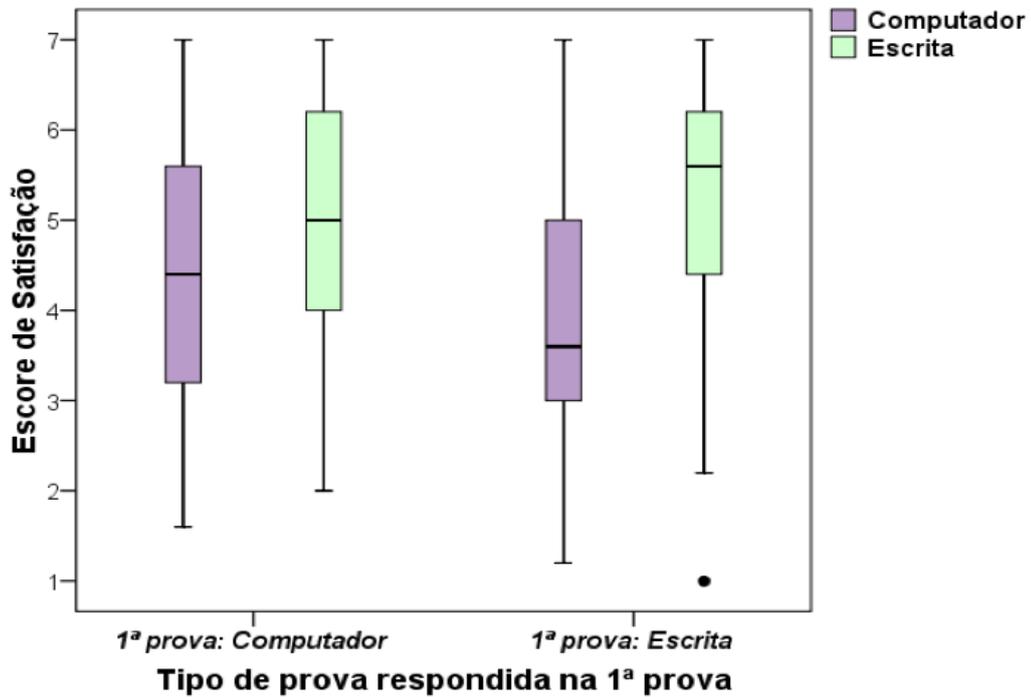
Fonte: Dados referentes ao estudo

Gráfico 3 - *Boxplot* dos escores de estresse referente à comparação entre os 2 tipos de provas considerando o tipo de prova respondida na 1ª fase.



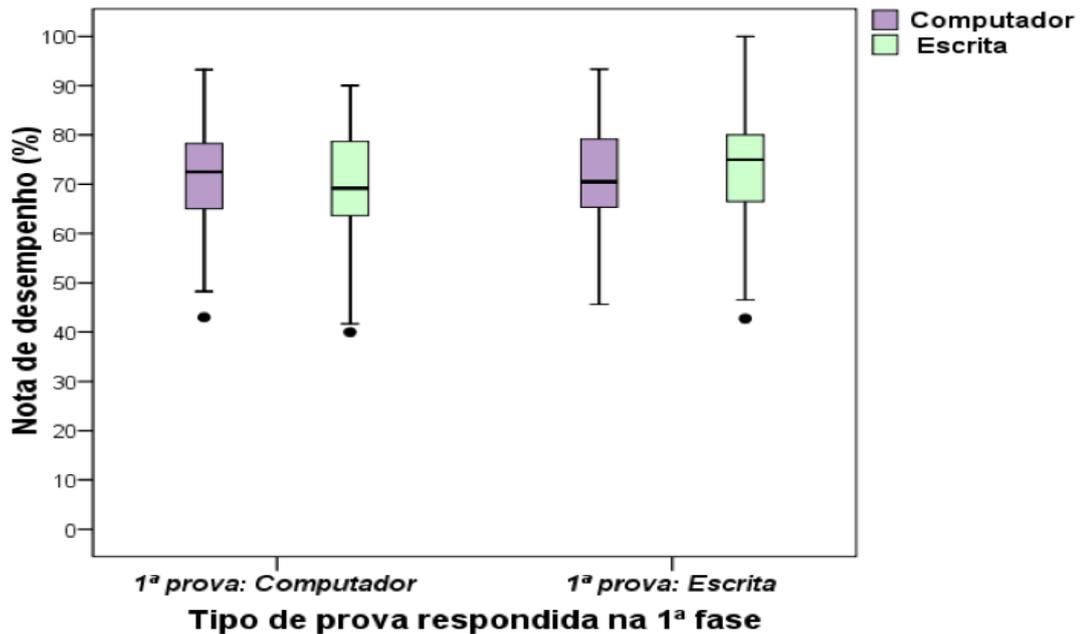
Fonte: Dados referentes ao estudo

Gráfico 4 - *Boxplot* dos Escores de Satisfação referente à comparação entre os 2 tipos de provas considerando o tipo de prova respondida na 1ª fase.



Fonte: Dados referentes ao estudo

Gráfico 5 - *Boxplot* das notas de desempenho dos alunos referente à comparação entre os 2 tipos de provas considerando o tipo de prova respondida na 1ª fase.



Fonte: Dados referentes ao estudo

A análise de correlação demonstrou fraca correlação entre os escores estresse e satisfação. Por sua vez, não foram observadas correlações entre os escores de estresse e satisfação e a nota de desempenho do aluno, independentemente da fase do estudo e do tipo de prova (TAB.5 e 6). Nota-se que a correlação entre os escores estresse e satisfação é negativa ($r < 0$), isto é, quanto maior a medida de um escore, menor a medida do outro escore.

Tabela 5 - Análise de Correlação entre Escore de Estresse, Escore de Satisfação e Nota de desempenho avaliada nas duas fases do estudo, por tipo de prova

Análise de correlação	n	r	p
<u>1ª fase do estudo</u>			
1ª prova: Computador			
<i>Estresse × Satisfação</i>	69	-0,39	0,001
<i>Estresse × Nota</i>	69	-0,13	0,304
<i>Satisfação × Nota</i>	69	0,22	0,070
1ª prova: Escrita			
<i>Estresse × Satisfação</i>	66	-0,16	0,192
<i>Estresse × Nota</i>	66	0,08	0,510
<i>Satisfação × Nota</i>	66	0,05	0,703
<u>2ª fase do estudo</u>			
2ª prova: Computador			
<i>Estresse × Satisfação</i>	69	-0,06	0,636
<i>Estresse × Nota</i>	69	-0,14	0,252
<i>Satisfação × Nota</i>	69	0,18	0,160
2ª prova: Escrita			
<i>Estresse × Satisfação</i>	66	-0,12	0,359
<i>Estresse × Nota</i>	66	-0,10	0,409
<i>Satisfação × Nota</i>	66	-0,12	0,343

Fonte: Dados referentes ao estudo

Tabela 6 - Análise de correlação entre escore de estresse, escore de satisfação e nota de desempenho avaliada independentemente da fase do estudo, por tipo de prova.

Análise de correlação	n	r	p
Prova no Computador			
<i>Estresse × Satisfação</i>	135	-0,23	0,008
<i>Estresse × Nota</i>	135	-0,13	0,126
<i>Satisfação × Nota</i>	135	0,20	0,020
Prova Escrita			
<i>Estresse × Satisfação</i>	135	-0,13	0,150
<i>Estresse × Nota</i>	135	-0,01	0,943
<i>Satisfação × Nota</i>	135	-0,03	0,745

Fonte: Dados referentes ao estudo

Com o objetivo de se avaliar a confiabilidade (consistência interna) das afirmativas que compõem cada um dos dois escores que avaliam o estresse e satisfação, com as provas aplicadas no computador ou de forma escrita (Convencional), a medida do Alfa de Cronbach foi utilizada. Os dois escores de interesse apresentaram medidas totais de Alfa de Cronbach superiores a 0,60, mostrando níveis de consistência interna satisfatórios, isto é, as questões/afirmativas conjuntamente estão medindo um mesmo Domínio / Dimensão (ou seja, na presente pesquisa, os escores) (TAB. 7).

Tabela 7 - Medida de confiabilidade da consistência interna das questões pertencentes aos escores de interesse – Prova no Computador.

(Continua)

Escore	Coefficiente de correlação Item-total	Alfa de Cronbach se questão retirada
• <i>Escore → Estresse</i>		
Q1	0,75	0,90
Q2	0,77	0,90
Q3	0,82	0,90
Q4	0,82	0,90
Q5	0,78	0,90
Q6	0,65	0,90
Q7	0,46	0,91
Q8	0,81	0,90
Q9	0,43	0,91
Q10	0,68	0,90

Tabela 7 - Medida de confiabilidade da consistência interna das questões pertencentes aos escores de interesse – Prova no Computador.

(Conclusão)		
Escore	Coeficiente de correlação Item-total	Alfa de <i>Cronbach</i> se questão retirada
Q11	0,71	0,90
<i>Q12</i>	0,66	0,90
Q13 INV	0,27	0,91
<i>Q14 INV</i>	0,15	0,92
Q15 INV	0,12	0,92
<i>Alfa de Cronbach Total → 0,91</i>		
<hr/>		
• Escore → Satisfação		
Q1	0,72	0,69
Q2	0,71	0,69
Q3 INV	0,39	0,80
<i>Q4</i>	0,21	0,84
Q5	0,85	0,65
<i>Alfa de Cronbach Total → 0,79</i>		

INV → A questão teve o sentido invertido, por exemplo. Q13 INV = (8 – Q13).

Fonte: Dados referentes ao estudo

Quanto à aplicação das provas de forma escrita aos alunos (TAB. 8), a análise é análoga, com apresentação de parâmetros não satisfatórios nas mesmas questões relativas à prova aplicada via computador. Este resultado mostra que independentemente do tipo de prova aplicado, a consistência interna dos escores de satisfação e estresse é semelhante. De um modo geral, os escores construídos revelaram-se satisfatórios quanto à consistência interna do conjunto de afirmativas presentes em cada escore.

Tabela 8 - Medida de confiabilidade da consistência interna das questões pertencentes aos escores de interesse – Prova Escrita.

Escore	Coefficiente de correlação Item-total	Alfa de Cronbach se questão retirada
• <i>Escore → Estresse</i>		
Q1	0,73	0,87
Q2	0,72	0,87
Q3	0,79	0,87
Q4	0,78	0,87
Q5	0,81	0,87
Q6	0,65	0,88
Q7	0,48	0,88
Q8	0,73	0,87
Q9	0,39	0,89
Q10	0,65	0,88
Q11	0,68	0,87
Q12	0,64	0,88
Q13 INV	0,28	0,89
Q14 INV	0,12	0,89
Q15 INV	-0,10	0,91
		<i>Alfa de Cronbach Total → 0,89</i>
• <i>Escore → Satisfação</i>		
Q1	0,71	0,65
Q2	0,71	0,65
Q3 INV	0,51	0,73
Q4	0,10	0,86
Q5	0,73	0,66
		<i>Alfa de Cronbach Total → 0,76</i>

Base de dados: 135 alunos

INV → A questão teve o sentido invertido, por exemplo. Q13 INV = (8 – Q13).

Fonte: Dados referentes ao estudo

As distribuições das respostas dos alunos para cada uma das afirmativas dos escores relacionados ao estresse e à satisfação, por tipo de prova aplicada, estão apresentadas no Apêndice B, visando a ampliar a compreensão acerca da distribuição das respostas em cada afirmativa, por escore e tipo de prova aplicada ao aluno.

6 DISCUSSÃO

Demonstrou-se neste estudo que o método de realização das provas não interferiu no desempenho dos acadêmicos, comparando-se as avaliações com uso da tecnologia ou provas convencionais escritas em papel. Educadores e avaliadores buscam assegurar que haja equivalência de resultados entre os métodos (LAURIE; BRIDGLALL; ARSENEAULT, 2015). De acordo com o entendimento desses autores, até os anos 90, havia diferença a favor das provas escritas, embora discreta, e que a construção da avaliação pode ser responsável por esse fator ser mais ou menos relevante. Deve-se ainda considerar que o treinamento anterior à avaliação melhora os resultados.

A aleatoriedade na constituição dos grupos avaliados garantiu a semelhança entre eles, uma vez que os dados sociodemográficos não interferiram nos resultados obtidos pelos acadêmicos envolvidos nas avaliações, seja idade, estado civil, cor, origem, período do curso, sexo, origem do ensino médio e ano de ingresso na faculdade. Clariana e Wallace (2002) obtiveram resultados semelhantes na avaliação sociodemográfica. Apenas em relação à etnia houve melhor resultado nos testes em computador, dos homens brancos em comparação com as outras raças.

Não foi verificada diferença entre o desempenho dos acadêmicos entre as provas, mesmo que não se tenha realizado treinamento no computador, prévio à aplicação das provas neste estudo. Clariana e Wallace (2002) citam trabalhos realizados em 1991, cujos resultados de testes demonstraram notas muito superiores em inglês e matemática em provas escritas, fato que não se repetiu em 2001, quando não foram observadas diferenças nos escores. Mesmo pequenas, as diferenças devem ser consideradas porque são importantes para os estudantes.

Laurie, Bridglall e Arseneault (2015) afirmam que, durante muito tempo, acreditava-se que o acesso do estudante à tecnologia, ou seja, seu conhecimento prévio de computação seria um fator facilitador para a execução das provas e para a transição das provas escritas para as que se baseiam em tecnologia. Com a facilidade dos *smartphones* na palma da mão, oferecendo acesso às mídias sociais a partir da internet, e também a pesquisas rápidas, essa geração digital possui facilidades no uso de tecnologia, diferentemente da geração anterior, que não estava familiarizada com tais recursos.

Ainda Laurie, Bridglall e Arseneault (2015) citam, em artigo, que os autores, ao descreverem o uso inicial do computador para execução de dissertações, afirmavam que os alunos, ao escreverem uma dissertação no papel, gastavam 50% menos tempo para executar a mesma tarefa que os colegas usando computadores. Cabe ressaltar, entretanto, que com a prática, essa diferença não se mantinha. No presente estudo, demonstrou-se que o fator tempo para leitura e execução das avaliações foi o mesmo entre os grupos das avaliações, usando-se computador ou prova escrita.

Em relação aos escores de estresse e satisfação, comparando prova escrita e prova no computador entre os acadêmicos do sétimo período e do décimo segundo período, não houve diferença entre os períodos do curso, em relação a esses escores. Não foi encontrada na literatura pesquisada comparação entre acadêmicos em períodos distintos, realizando avaliações com conteúdos idênticos, uma escrita e outra no computador. Esperava-se que os concluintes apresentassem um escore de stress menor que os alunos que finalizaram o ciclo básico, mas isso não ocorreu.

Não se encontrou, nos dados avaliados, diferença em relação ao estresse, seja por idade ou por gênero, apenas pelo uso da tecnologia em si. Hassanien et al. (2013) também afirmam que o estresse com o uso da tecnologia não altera o desempenho nas avaliações em estudantes de Medicina.

Os acadêmicos envolvidos neste estudo apresentaram maior estresse na prova realizada no computador, na primeira fase do estudo. Em relação à satisfação, comparando-se as provas no computador e escritas, houve maior satisfação na realização da prova escrita, nas duas fases do estudo.

Ao relacionarem-se os escores entre o desempenho, estresse e satisfação, observa-se que, quando comparados os acadêmicos randomizados para a primeira avaliação no computador e a segunda prova escrita, existiu uma diferença significativa em relação ao estresse, sendo que a avaliação realizada no computador causou maior estresse do que a prova escrita, e o escore de satisfação foi superior com a prova escrita; no entanto, esses dados não interferiram no desempenho acadêmico.

Quando a primeira prova foi escrita, e a segunda prova realizada no computador, o escore de estresse foi semelhante entre as avaliações. Inferindo-se que o contato e troca de experiências entre os alunos favoreceram esse resultado. A experiência com o método altera os resultados e nivela o estresse por ser prova, independentemente do formato. Em relação à satisfação manteve-se a preferência por provas escritas, ainda sem diferença entre os escores de desempenho. O que não confirma os resultados de Appel et al., (2005), em que os testes no computador foram bem aceitos por 63,65% a 67,69 % dos acadêmicos avaliados por eles.

Após avaliação dos resultados ofereceu-se aos alunos do semestre seguinte um pré-teste com dez questões, em sala de aula, para que os acadêmicos se familiarizassem com o novo método, com resultados, até o momento, superiores em relação à satisfação.

Ao analisar a correlação entre os escores de estresse, de satisfação e desempenho acadêmico, nas duas fases do estudo, por tipo de prova, pode-se afirmar que na primeira fase do estudo, prova no computador, comparando-se estresse e satisfação, encontra-se uma relação inversa, isto é, quanto maior o estresse, menor a satisfação. Na correlação entre estresse e nota, percebeu-se que o estresse não interferiu na nota. Nem mesmo a menor satisfação com a avaliação no computador interferiu no desempenho acadêmico.

Cantillon, Irish e Sales (2004) sugerem que o treinamento prévio melhora os resultados e diminui o estresse. Interessante observar que, após 14 anos da publicação desse artigo, a familiaridade com a tecnologia não se mostrou suficiente para a adaptação aos novos testes usando tecnologia. Esse fato não se provou verdadeiro no estudo em questão. Hassanien et al. (2013) orientam aos que pedem sugestões para melhorar as avaliações feitas no computador, com a afirmativa: “ofereça mais provas-piloto antes do exame final”.

Sugerir mudanças no modo de avaliar traz consequências para docentes e discentes e, embora os artigos pesquisados mostrem que o mundo tem utilizado o computador como ferramenta agregadora, há pelo menos duas décadas, não se tinha conhecimento do método, seus desafios e dificuldades na implantação.

Modificar o que já funciona, alterando a visão do modo de avaliar de todo um departamento envolve persuasão e perseverança. Demonstrar que, ao usar a tecnologia, não se prejudica o resultado acadêmico, com diminuição das notas e com efeito negativo no desempenho dos

estudantes, foi o primeiro ponto a ser explicado ao departamento de Ginecologia e Obstetrícia. Alertou-se para o fato de que, mesmo que haja aumento do trabalho para se adaptar às questões de provas em um formato padronizado e mais atraente, com uso de figuras e ou gráficos, quando possível, o ganho futuro seria compensador, ter-se-ia menos trabalho na elaboração de novas questões e novas avaliações, lançando-se mão do Banco já pronto. Embora esse processo de elaborar questões seja contínuo, a prática o torna mais fácil. Bull (1999) afirma que formar o banco de questões é fator imprescindível e de preocupação nesse processo. Deve-se construí-lo com questões de qualidade e confiabilidade cujo conteúdo satisfaça a todas as partes envolvidas, discentes e docentes.

Seguiram-se as sugestões de Chalmers e Mcausland (2002) e Hassnien (2013) para implantação da avaliação no computador, definindo o objetivo pedagógico; em curto prazo, melhorar a qualidade das questões e oferecer as notas e o feedback imediatos ao término da avaliação, com o intuito de reforçar com a correção, a aprendizagem e a fixação do conteúdo.

Em médio prazo, o objetivo será usar a tecnologia também nas avaliações formativas, o que demandará comprometimento ainda maior de todos.

Não se tinha experiência de usar a plataforma Moodle, com a finalidade de execução de uma avaliação, por isso fez-se um pré-teste a fim de verificar a usabilidade da plataforma. Logo no início, percebeu-se a necessidade de que as questões fossem embaralhadas, porque alunos comunicam entre si, e o laboratório de informática possui monitores sem divisória apropriada (FIG. 2), uma vez que os equipamentos não foram dispostos com o objetivo de ser utilizado para avaliações. Assim, a preocupação adicional foi com respeito à “cola” entre os alunos. (FIG. 1). Essa preocupação com a comunicação entre os alunos também é relatada por Chalmers e Mcausland (2002) e Canole (2005), ao lado de cuidados com a segurança dos dados coletados e enviados para a secretaria e banco de dados (SIM; HOLIFIELD; BROWN, 2004; CANOLE, 2005).

Percebeu-se ao término da coleta dos dados, que o número de questões ideais para uma avaliação, em uma hora, seria entre 40 a 50 questões, dessa forma haverá menos tempo para a tentativa de comunicação entre os alunos. Cantillon, Irish e Sales (2004) relatam que a avaliação para licença médica nos Estados Unidos utiliza 50 questões por hora. Hassanien et al. (2013) usaram em sua pesquisa com acadêmicos de Medicina, 40 questões por hora, 80 questões para

serem respondidas em 120 minutos. No caso do presente estudo, o tempo mostrou-se suficiente, e nenhum acadêmico deixou de terminar a avaliação.

Pela disponibilidade prévia de uma plataforma com finalidade avaliativa, neste estudo não se compartilham as preocupações relatadas por Cantillon, Irish e Sales (2004), Bull (1999) e Chalmers e McAusland (2002) sobre a implantação de um *software*, muitas vezes dispendioso.

Várias foram as dificuldades encontradas ao iniciar-se este estudo, como a falta do conhecimento dos termos internacionais utilizados para avaliações no computador, ou seja, desconhecimento das palavras-chave corretas que só foram percebidas no terço final da dissertação, mas que permitiram dessa forma direcionar a pesquisa. Por outro lado, ter conhecimento da existência de uma conferência específica englobando as tecnologias envolvidas no processo de avaliação alterou definitivamente a percepção do impacto que seu uso exerce no processo ensino/aprendizagem, no cenário da educação atual e permitiu embasamento para persistir nessa linha de pesquisa.

Como limitações deste estudo, mencionam-se os seguintes itens: a falta de treinamento dos alunos com o método de prova usando o computador anteriormente às avaliações; a não-preparação para enfrentar casos fora da normalidade esperada, como por exemplo, o fato de uma acadêmica precisar de mais tempo para resolver a prova por ser portadora de dislexia; escassez de tempo para a coleta de dados e para sua interpretação.

É praticamente impossível realizar um estudo livre de limitações, no entanto antever possíveis problemas e traçar estratégias para enfrentá-los deve fazer parte de todo e qualquer desenho de investigação. Após o término da coleta de dados, percebeu-se, em uma pesquisa bibliográfica final, a existência de um questionário específico para avaliar o estresse ao utilizarem-se tecnologias, poder-se-ia, assim, direcionar os resultados nesse quesito.

7 CONCLUSÃO

Neste estudo avaliaram-se duas turmas distintas, do curso de Medicina da UNIFENAS, *campus* de Alfenas, com o intuito de oferecer uma forma alternativa às avaliações convencionais. Ministrou-se uma avaliação em papel e outra no computador para cada aluno cursando Ginecologia e Obstetrícia, no primeiro semestre de 2018. Não foram encontradas diferenças nos escores de desempenho acadêmico entre os dois processos.

A falta de treinamento prévio e familiaridade dos estudantes com o método parece ter sido o fator responsável pela preferência da maioria pelas provas escritas.

Professores que queiram implementar este tipo de avaliação devem introduzir a metodologia ao longo do curso, com testes-piloto, por exemplo.

REFERÊNCIAS

- APPEL, J. et al. Educational assessment guidelines: a clerkship directors in internal medicine commentary. **The American Journal of Medicine**, [S.l.], v. 113, n. 2, p. 172-179, 2002.
- AZIMI, M. The Relationship between anxiety and Test-Taking C-Test and Cloze-Test. **Malaysian Online Journal of Educational Sciences**, [S.l.], v. 4, n. 1, p. 30-42, 2016.
- BORGES, M. C. et al. Avaliação formativa e feedback como ferramenta de aprendizado na formação de profissionais da saúde. **Medicina**, Ribeirão Preto, v. 47, n. 3, p. 324-331, 2014. Disponível em: http://revista.fmrp.usp.br/2014/vol47n3/11_Avaliacao-formativa-e-feedback-como-ferramenta-de-aprendizado-na-formacao-de-profissionais-da-saude.pdf. Acesso em: 20 out. 2018
- BRASIL, Ministério do Planejamento e orçamento: Fundação Instituto Brasileiro de Geografia e Estatística. **500 anos de povoamento**. Rio de Janeiro: IBGE, 2000. Disponível em: <https://biblioteca.ibge.gov.br/visualizacao/livros/liv6687.pdf>. Acesso em: 20 out 2018.
- BULL, J. Computer-assisted assessment: Impact on higher education institutions. **Educational Technology & Society**, Luton, United Kingdom, v. 2, n. 3, p. 123-126, 1999.
- BULL, J.; DANSON, M. **Computer-assisted assessment (CAA)**. New York, LTSN, 2004.
- CANOLE, G.; WARBURTON, B. A review of computer-assisted assessment. **ALT-J**, [S.l.], v. 13, n. 1, p. 17-31, 2005.
- CANTILLON, P.; IRISH, B.; SALES, D. Using computers for assessment in medicine. **BMJ**, London, v. 329, n. 7466, p. 606-609, 2004.
- CHALMERS, D.; MCAUSLAND, W. D. M. Computer-assisted assessment. **The handbook for economics lecturers: Assessment**, [S.l.], p. 2-20, 2002.
- CLARIANA, R.; WALLACE, P. Paper-based versus computer-based assessment: key factors associated with the test mode effect. **British Journal of Educational Technology**, online, v. 33, n. 5, p. 593-602, 2002. Disponível em: <https://www.semanticscholar.org/paper/Paper-based-versus-computer-based-assessment%3A-key-Clariana-Wallace/3e8aff00cb390786995d0abb473d8233866b3f9e>. Acesso em: 20 out 2018.
- COLLINS, A.; HALVERSON, R. The second educational revolution: Rethinking education in the age of technology. **Journal of computer assisted learning**, [S.l.], v. 26, n. 1, p. 18-27, 2010.
- EPSTEIN, R. M. Assessment in medical education. **New England Journal of Medicine**, [S.l.], v. 356, n. 4, p. 387-396, 2007.

FERREIRA, A. R. L.; WEYH, C. B. Tecnologia da informação no processo da avaliação como estratégia para qualificar a prática docente. **CIET: EnPED**, 2018.

HARGREAVES, M. et al. Computer or paper? That is the question: does the medium in which assessment questions are presented affect children's performance in mathematics? **Educational Research**, v. 46, n. 1, p. 29-42, 2004.

HASSANIEN, M. A. et al. A six step approach for developing computer based assessment in medical education. **Medical teacher**, Dundee - United Kingdom, v. 35, n. sup1, p. S15-S19, 2013.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA, Censo 2010. Disponível em: <https://censo2010.ibge.gov.br/>. Acesso em: 20 out 2018.

LAURIE, R.; BRIDGLALL, B. L.; ARSENEAULT, P. Investigating the effect of computer-administered versus traditional paper and pencil Assessments on Student Writing Achievement. **SAGE Open**, [S.l.], v. 5, n. 2, p. 215, 2015.

LIKERT, R. A technique for the measurement of attitudes. **Archives of Psychology**, [S.l.], v. 22, n. 140, p. 1-55, 1932.

LIMA, R. L. et al. Stress among medical students and academic performance. **Revista Brasileira de Educação Médica**, Brasília, v. 40, n. 4, p. 678-684, 2016.

LIPP, M, E. N.; **Manual do inventário de sintomas de stress para adultos de Lipp (ISSL)**. São Paulo, Casa do Psicólogo, v. 76, 2000.

LUECHT, R. M. et al. Testing the test: a comparative study of a Comprehensive Basic Sciences Test Using Paper-and-pencil and Computerized Formats. **Academic Medicine**, United States, v. 73, n. 10, p. S51-53, 1998.

OGILVIE, R. W.; TRUSK, T. C.; BLUE, A. V. Students' attitudes towards computer testing in a basic science course. **Medical education**, [S.l.], v. 33, n. 11, p. 828-831, 1999

PANÚNCIO-PINTO, M. P.; TRONCON, L. E. A; Avaliação do estudante–aspectos gerais. **Medicina**. Ribeirao Preto, v. 47, n. 3, p. 314-323, 2014. Disponível em: http://revista.fmrp.usp.br/2014/vol47n3/10_valiacao-do-estudante-aspectos-gerais.pdf. Acesso em: 20 nov 2018

SIM, G.; HOLIFIELD, P.; BROWN, M. Implementation of computer assisted assessment: lessons from the literature. **ALT-J**, online, v. 12, n. 3, p. 215-229, 2004. Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/0968776042000259546>. Acesso em: 20 nov 2018.

SPSS 17.0 for Windows (Software Estatístico).

WIJERATHNE, B.; RATHNAYAKE, G. Medical students' attitudes and perspectives regarding novel computer-based practical spot tests compared to traditional practical spot tests. **Canadian medical education journal**, [S.l.]. v. 4, n. 2, p. 41, 2013. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4563600/>. Acesso em: 20 nov 2018.

YATES, J. When did they leave, and why? A retrospective case study of attrition on the Nottingham undergraduate medical course. **BMC medical education**, [S.l.], v. 12, n. 1, p. 43, 2012.

APÊNDICE A -Questionário da pesquisa – Fase 1

Pág. 1 / 2

Este questionário consta de duas partes. Por favor, responda todas as perguntas c/ sinceridade. Os questionários serão avaliados **SOMENTE** de forma conjunta, portanto, as resposta individuais não serão divulgadas uma vez que estarão sob sigilo.

Por favor, evite deixar questões sem resposta.

DADOS SOCIO-DEMOGRÁFICOS E COMPORTAMENTAIS

NOME:

Nº de matrícula:

1. Quantos anos completos você tem?

2. Qual o seu sexo? 1.Masculino 2.Feminino

3. Qual o seu estado civil? 1.Solteiro(a) 3.Separado(a)
 2.Casado(a) / União estável 4.Viúvo(a)

4. Qual sua etnia? 1.Branca 3.Negra 5.Indígena
 2.Parda 4.Amarela 6.Outra

5. Qual a origem do ensino médio? 1.Público 3.Público e Privado
 2.Privado 4.Outra

6. Qual o ano de ingresso na faculdade?

7. Você realiza algum trabalho remunerado? 1.Sim (Vá p/ Q8) 2.Não (Vá p/ Q9)

8. Se "SIM", Quantas horas trabalhadas, em média, por semana?

SATISFAÇÃO COM AS PROVAS: Computador versus Escrita Convencional

Nas questões a seguir, **MARQUE COM UM X** o número que **MELHOR** corresponde à sua opinião. Sua resposta deverá expressar a sua opinião pessoal sobre o afirmativa proposta. **NÃO HAVENDO, PORTANTO, RESPOSTAS CERTAS OU ERRADAS.**

RESPONDA CONFORME ESCALA ABAIXO: (NÃO DEIXE NENHUM ITEM SEM RESPOSTA)

Discordo **1** **2** **3** **4** **5** **6** **7** Concordo
Totalmente Totalmente

RESPONDA AS AFIRMATIVAS ABAIXO:

	1	2	3	4	5	6	7
16. Considero que o formato da prova no computador é mais adequado.	<input type="checkbox"/>						
17. O layout da prova no computador é mais atraente, facilitando o entendimento e interpretação das questões.	<input type="checkbox"/>						
18. A realização de prova no computador é desconfortável e piora meu desempenho.	<input type="checkbox"/>						
19. O tempo para realização da prova foi suficiente para o término da mesma.	<input type="checkbox"/>						
20. A realização de prova no computador me deixa satisfeito.	<input type="checkbox"/>						

**AGRADEÇO SUA
IMPORTANTE COLABORAÇÃO**

SATISFAÇÃO COM AS PROVAS: Computador versus Escrita Convencional

Nas questões a seguir, **MARQUE COM UM X** o número que **MELHOR** corresponde à sua opinião. Sua resposta deverá expressar a sua opinião pessoal sobre o afirmativa proposta. **NÃO HAVENDO, PORTANTO, RESPOSTAS CERTAS OU ERRADAS.**

RESPONDA CONFORME ESCALA ABAIXO: (NÃO DEIXE NENHUM ITEM SEM RESPOSTA)

Discordo	1	2	3	4	5	6	7	Concordo
Totalmente	<input type="checkbox"/>	Totalmente						

RESPONDA AS AFIRMATIVAS ABAIXO:

	1	2	3	4	5	6	7
16. Considero que o formato da prova no computador é mais adequado.	<input type="checkbox"/>						
17. O layout da prova no computador é mais atraente, facilitando o entendimento e interpretação das questões.	<input type="checkbox"/>						
18. A realização de prova no computador é desconfortável e piora meu desempenho.	<input type="checkbox"/>						
19. O tempo para realização da prova foi suficiente para o término da mesma.	<input type="checkbox"/>						
20. A realização de prova no computador me deixa satisfeito.	<input type="checkbox"/>						

<p>AGRADEÇO SUA</p> <p>IMPORTANTE COLABORAÇÃO</p>

SEU SENTIMENTO EM REALIZAR PROVAS

Nas questões a seguir, MARQUE COM UM X o número que MELHOR corresponde à sua opinião. Sua resposta deverá expressar a sua opinião pessoal sobre o afirmativa proposta. NÃO HAVENDO, PORTANTO, RESPOSTAS CERTAS OU ERRADAS.

RESPONDA CONFORME ESCALA ABAIXO: (NÃO DEIXE NENHUM ITEM SEM RESPOSTA)

Discordo **1** **2** **3** **4** **5** **6** **7** Concordo
Totalmente Totalmente

RESPONDA A AS AFIRMATIVAS ABAIXO:

	1	2	3	4	5	6	7
6. Antes das provas, geralmente, sinto frio na barriga.	<input type="checkbox"/>						
7. Antes das provas, geralmente, meu coração bate mais rápido.	<input type="checkbox"/>						
8. Antes das provas, geralmente, me sinto apreensivo.	<input type="checkbox"/>						
9. Antes das provas, geralmente, me sinto angustiado.	<input type="checkbox"/>						
10. Antes das provas, geralmente, me sinto pressionado.	<input type="checkbox"/>						
11. Antes das provas, geralmente, minha boca seca.	<input type="checkbox"/>						
12. Antes das provas, geralmente, preciso ir ao banheiro mais vezes.	<input type="checkbox"/>						
13. Antes das provas, geralmente, me sinto estressado.	<input type="checkbox"/>						
14. Antes das provas, geralmente, me sinto mais estressado com a prova escrita.	<input type="checkbox"/>						
15. Na realização da prova fico me perguntando se meu desempenho será suficiente.	<input type="checkbox"/>						
16. Na realização da prova fico preocupado com a possibilidade de reprovação.	<input type="checkbox"/>						
17. Na realização da prova me preocupo com a quantidade de conteúdo cobrado.	<input type="checkbox"/>						
18. Na realização da prova confio no meu desempenho.	<input type="checkbox"/>						
19. Na realização da prova fico satisfeito comigo mesmo.	<input type="checkbox"/>						
20. Na realização da prova me sinto mais confiante na prova escrita.	<input type="checkbox"/>						

**AGRADEÇO SUA
IMPORTANTE COLABORAÇÃO**

APÊNDICE C -Termo de Consentimento Livre e Esclarecimento

Você está sendo convidado para participar como voluntário de uma pesquisa proposta pela Universidade José do Rosário Velano, descrita em detalhes a seguir.

Para decidir se você deve concordar ou não em participar desta pesquisa, leia atentamente todos os itens a seguir que irão lhe informar e esclarecer todos os procedimentos, riscos e benefícios aos quais você deve seguir, segundo as exigências da Resolução 510/2016 do Conselho Nacional de Saúde do Ministério da Saúde.

1. Identificação do(a) voluntário(a) da pesquisa:

Nome:

Gênero:

Identidade:

Órgão Expedidor:

Data de Nascimento: / /

Responsável Legal (se aplicável):

Gênero:

Identidade:

Órgão Expedidor:

Data de Nascimento: / /

2. Dados da pesquisa:

a. Título do Projeto: Avaliação do desempenho e do grau de satisfação dos alunos de Ginecologia e Obstetrícia G.O comparando testes escritos convencionais e testes no computador.

b. Universidade/Departamento/Faculdade/Curso: Universidade José do Rosário Velano . Departamento de Ginecologia e Obstetrícia Faculdade de ciências Médicas, Medicina Humana

c. Projeto: (X) Unicêntrico () Multicêntrico

d. Instituição Co-participante:

e. Patrocinador:

f. Professor Orientador: Dra Maria Aparecida Turci

Pesquisador Responsável: () Estudante de Pós-graduação () Professor Orientador

3 Objetivo da pesquisa:

Avaliar o estresse, a satisfação e o desempenho dos alunos do sétimo e décimo segundo períodos, em Ginecologia e Obstetrícia, comparando as provas escritas convencionais e as provas realizadas em computador

5. Justificativa da pesquisa

Parece que existem vantagens dos testes *online*/ computador sobre os testes escritos tradicionais. Entre os benefícios das provas em ambiente virtual, tem-se a possibilidade de fornecer correção em tempo real e a realização de testes a distância. No entanto, a larga implementação dos testes *online* depende da comparação entre os dois tipos de avaliação em relação ao estresse gerado pela prova, desempenho e grau de satisfação entre os estudantes que os realizam. Existem ainda lacunas em relação a essas questões. Portanto, o objetivo do estudo proposto é avaliar essas questões nos alunos em Ginecologia e Obstetrícia do sétimo e décimo segundo períodos, comparando as provas escritas convencionais e as provas realizadas no computador.

6. Descrição detalhada e explicação dos procedimentos realizados:

Os alunos do sétimo e décimo segundo períodos serão submetidos a uma prova teste convencional, ea uma prova em computador, na qual serão avaliados a satisfação e o desempenho dos alunos em ambos os formatos.

7. Descrição dos desconfortos e riscos da pesquisa:

(x) Risco Mínimo () Risco Baixo () Risco Médio () Risco Alto

Pode-se sentir o desconforto e possível inibição em responder às perguntas do questionário. Não há riscos previsíveis.

8. Descrição dos benefícios da pesquisa:

Contribuir para gerar informações sobre os dois diferentes tipos de provas e aperfeiçoamento das avaliações, como testes a distância, correção online ou presencial e feedback imediatos. Permitir que a Universidade possa armazenar os dados acadêmicos em espaço virtual.

9. Despesas, compensações e indenizações:

- a. Você não terá despesa pessoal nessa pesquisa incluindo transporte, exames e consultas.
- b. Você não terá compensação financeira relacionada a sua participação nessa pesquisa.

10. Direito de confidencialidade:

- a. Você tem assegurado que todas as suas informações pessoais obtidas durante a pesquisa serão consideradas estritamente confidenciais e os registros estarão disponíveis apenas para os pesquisadores envolvidos no estudo.
- b. Os resultados obtidos nessa pesquisa poderão ser publicados com fins científicos, mas sua identidade será mantida em sigilo.
- c. Imagens ou fotografias que possam ser realizadas, se forem publicadas, não permitirão sua identificação.

11. Acesso aos resultados da pesquisa:

Você tem direito de acesso atualizado aos resultados da pesquisa.

12. Liberdade de retirada do consentimento:

Você tem direito de retirar seu consentimento, a qualquer momento, deixando de participar da pesquisa, sem qualquer prejuízo dentro da instituição.

13. Acesso aos profissionais responsáveis pela pesquisa:

Você tem garantido o acesso, em qualquer etapa da pesquisa, aos profissionais responsáveis por ela, para esclarecimento de eventuais dúvidas acerca de procedimentos, riscos, benefícios, etc., por meio do seguinte contato:

Professor Orientador: Maria Aparecida Turci Telefone: (31) 34974300

Email: maria.turci@unifenas.br

14. Acesso à instituição responsável pela pesquisa:

Você tem garantido o acesso, em qualquer etapa da pesquisa, à instituição responsável pelo estudo, para esclarecimento de eventuais dúvidas acerca dos procedimentos éticos, por meio do seguinte contato:

Comitê de Ética - UNIFENAS:

Rodovia MG 179, Km 0, Alfenas – MG Tel: (35) 3299-3137

Email: comitedeetica@unifenas.br

Horário: segunda à sexta-feira das 14:00h às 16:00h

APÊNDICE D -Termo de adesão

Declaro que fui informado verbalmente e por escrito sobre os dados desta pesquisa e minhas dúvidas com relação a minha participação foram satisfatoriamente respondidas.

Ficaram claros para mim quais são os propósitos do estudo, os procedimentos a serem realizados, os desconfortos e riscos, as garantias de confidencialidade e de esclarecimentos permanentes. Estou ciente também de que a minha participação é isenta de despesas e que tenho garantia do acesso aos pesquisadores e à instituição de ensino.

Tive tempo suficiente para decidir sobre minha participação e concordo voluntariamente em participar desta pesquisa e poderei retirar o meu consentimento a qualquer hora, antes ou durante sua realização, sem penalidades, prejuízo ou perda de qualquer benefício que eu possa ter adquirido.

A minha assinatura neste Termo de Consentimento Livre e Esclarecido dará autorização aos pesquisadores, ao patrocinador do estudo e ao Comitê de Ética em Pesquisa da Universidade José do Rosário Vellano, de utilizarem os dados obtidos quando se fizer necessário, incluindo sua divulgação, sempre preservando minha identidade.

Assino o presente documento em duas vias de igual teor e forma, ficando uma em minha posse.

Alfenas, de de

Assinatura Dactiloscópica

Voluntário

Representante Legal

Pesquisador Responsável Voluntário Representante Legal

APÊNDICE E -Aprovação pelo Comitê de Ética da Unifenas

UNIVERSIDADE JOSÉ
ROSÁRIO VELLANO/UNIFENAS



PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: COMPARAÇÃO ENTRE TESTES ESCRITOS CONVENCIONAIS E TESTES REALIZADOS NO COMPUTADOR: ESTRESSE, SATISFAÇÃO E DESEMPENHO DOS ALUNOS DO SÉTIMO E DÉCIMO SEGUNDO PERÍODOS DE MEDICINA DA UNIFENAS - ALFENAS, CURSANDO GINECOLOGIA E OBSTETRÍCIA.

Pesquisador: MERILZA RODRIGUES

Área Temática:

Versão: 1

CAAE: 84230818.5.0000.5143

Instituição Proponente: Universidade José Rosário Vellano/UNIFENAS

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 2.535.778

Apresentação do Projeto:

Adequada.

Objetivo da Pesquisa:

Adequado..

Avaliação dos Riscos e Benefícios:

Adequado.

Comentários e Considerações sobre a Pesquisa:

Pesquisa relevante.

Considerações sobre os Termos de apresentação obrigatória:

Adequados.

Conclusões ou Pendências e Lista de Inadequações:

Nada digno de nota.

Considerações Finais a critério do CEP:

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Endereço: Rodovia MG 179 km 0
 Bairro: Campus Universitário CEP: 37.130-000
 UF: MG Município: ALFENAS
 Telefone: (35)3299-3137 Fax: (35)3299-3137 E-mail: comiteetica@unifenas.br

UNIVERSIDADE JOSÉ
ROSÁRIO VELLANO/UNIFENAS



Continuação do Parecer: 2.535.778

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_1082499.pdf	04/03/2018 19:53:37		Aceito
Declaração de Instituição e Infraestrutura	declamerilza.pdf	04/03/2018 19:48:47	MERILZA RODRIGUES	Aceito
Projeto Detalhado / Brochura Investigador	Projeto_detalhado.docx	04/03/2018 19:47:18	MERILZA RODRIGUES	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLE.docx	04/03/2018 19:46:49	MERILZA RODRIGUES	Aceito
Folha de Rosto	folhaDeRosto.pdf	04/03/2018 19:46:16	MERILZA RODRIGUES	Aceito
Outros	Questionario_Fase2.docx	25/02/2018 12:08:19	MERILZA RODRIGUES	Aceito
Outros	Questionario_Fase1.docx	25/02/2018 12:07:59	MERILZA RODRIGUES	Aceito

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

ALFENAS, 09 de Março de 2018

Assinado por:
MARCELO REIS DA COSTA
(Coordenador)



Laboratório de informática, CEP (centro. De estudos e pesquisas HUAV)