Avaliação da eficiência de jogo educativo para o ensino do comportamento de distinguir e-mails legítimos de tentativas de *phishing*

Jasson Marques Fontoura Júnior[1], Marcelo Henrique Oliveira Henklain[1], Felipe Leite Lobo[1], Eduardo Luzeiro Feitosa[2]

[1]Departamento de Ciência da Computação - Universidade Federal de Roraima – Boa Vista – Roraima – Brasil

[2]Instituto de Ciência da Computação – Universidade Federal do Amazonas – Manaus – Amazonas – Brasil

jassonjr5@gmail.com, {marcelo.henklain, felipe.lobo}@ufrr.br, efeitosa@icomp.ufam.edu.br

Resumo. *O objetivo deste estudo foi desenvolver e avaliar a eficiência do jogo educativo Alerta no ensino do comportamento de "distinguir e-mails legítimos de tentativas de phishing". Na primeira iteração de desenvolvimento, participaram 4 estudantes de computação e, da segunda, 27 pessoas de áreas diversas. Os resultados foram promissores, evidenciando aprendizado do comportamento de distinção entre e-mail legítimo e com phishing, melhora de autoconfiança e indicadores positivos de usabilidade e satisfação. Estudos futuros precisam continuar aperfeiçoando e testando o jogo com uma amostra maior e mais diversa e, ainda, incluir novos objetivos de aprendizagem que capacitem as pessoas a lidarem com técnicas modernas de phishing.*

Abstract. *The aim of this study was to develop and evaluate the efficiency of the educational game Alerta for teaching the behavior of "distinguishing legitimate emails from phishing attempts". In the first development iteration, 4 computer science students participated, and in the second, 27 individuals from diverse fields. The results were promising, demonstrating learning in distinguishing between legitimate and phishing emails, improved self-confidence, and positive indicators of usability and satisfaction. Future studies need to continue refining and testing the game with a larger and more diverse sample, as well as incorporating new learning objectives to empower individuals to handle modern phishing techniques.*

1. Introdução

Segundo relatórios da APWG (2023, 2024), os ataques de *phishing* aumentaram 40% no mundo, chegando a mais de 960 mil no primeiro trimestre de 2024. Só no Brasil, existem mais de 5 mil páginas falsas detectadas, que são utilizadas para *phishing* [CERT.br, 2024]. Trata-se, portanto, de ameaça à segurança digital de indivíduos e organizações. Dentre os diversos vetores para tais ataques, e-mails persistem tendo grande efetividade entre os usuários e, por isso, merecem atenção [Souza et al., 2023].

Cientistas e profissionais de cibersegurança destacam a urgência de medidas educativas para mitigar essa ameaça porque a capacidade do usuário de identificar

phishing é crucial para impedi-lo [Rahman et al., 2020]. Campanhas de conscientização e métodos de ensino baseados em exposição de informações têm se mostrado pouco efetivos, uma vez que usuários persistem apresentando condutas inseguras e sofrendo as suas consequências [Guilherme et al., 2021; Kovačević et al., 2020; Wash, 2020].

É possível que a baixa efetividade ocorra porque os usuários são expostos apenas a regras, sem treino explícito [Alanazi et al., 2022.] na tarefa de reconhecer *phishing*. Jogos educativos podem ser promissores para esse tipo de treino, pois geram motivação mesmo em tarefas repetitivas, permitem a programação de reforçamento imediato para seleção dos comportamentos definidos como alvos do ensino e podem ser utilizados por muitas pessoas simultaneamente [Arachchilage et al., 2011; Farias et al., 2019; Gris et al., 2018]. Por essas razões, o objetivo deste estudo foi desenvolver e avaliar a eficiência de jogo educativo para ensinar o comportamento de "distinguir e-mails legítimos de tentativas de *phishing*". Após desenvolver duas iterações do jogo, buscamos responder às seguintes perguntas de pesquisa: **PP01.** O jogo foi eficiente na promoção de aprendizagens para todos os jogadores? **PP02.** Os jogadores aumentaram o seu grau de autoconfiança em relação à identificação de e-mails legítimos e tentativas de *phishing* após o jogo? **PP03.** Em que medida a experiência do jogo foi agradável? **PP04.** Qual o grau de usabilidade do jogo? Este trabalho possui seções de fundamentação teórica, trabalhos relacionados, método, resultados e discussão e conclusão.

2. Fundamentação Teórica

2.1. Ataques de *phishing*: Uma das principais ameaças do ciberespaço

Conforme CERT.br (2024), ataques de *phishing* são frequentes no Brasil, representando um problema grave para a economia e bem-estar das pessoas. Apesar disso e das campanhas de conscientização, usuários persistem emitindo condutas inseguras no uso da Internet, levando ao sucesso dos ciberataques [Guilherme et al., 2021].

Ataques de *phishing*, tipicamente, são realizados por e-mail e envolvem o uso de engenharia social para obter informações valiosas ou convencer o usuário a clicar em *link* ou arquivo anexo [Syafitri et al., 2022], ação que pode resultar na infecção do computador com *malwares* [Melo et al., 2011; Diorio et al., 2018] ou levar a sites falsos, nos quais o usuário é induzido a fornecer dados ou comprar algo. Segundo Scarfone et al. (2008), no e-mail, o atacante usa informações falsas e se passa por empresa ou pessoa conhecida, para atrair a atenção e confiança do usuário, podendo adotar técnicas de e-mail *spoofing*, para falsificar o cabeçalho do e-mail, fazendo-o parecer legítimo [Morgado et al., 2023].

Ao estudar sobre *phishing*, identificamos várias pesquisas que se concentram na intervenção sobre comportamentos de cibersegurança [e.g., Hartwig et al., 2021], dada a sua importância estratégica para a sociedade [Ruoslahti et al., 2021]. São reconhecidas como promissoras as intervenções baseadas em recursos tecnológicos, pois viabilizam treinamento de pessoas em larga escala. Neste estudo, trabalharemos nesta direção. Para tanto, precisamos compreender o que é e como intervir sobre o comportamento humano.

2.2. Teoria sobre os processos de ensinar e aprender

Para a teoria analítico-comportamental, o comportamento é um sistema de interações entre ambiente antecedente, ações de uma pessoa e ambiente subsequente [Skinner, 1981]. "distinguir e-mails legítimos de tentativas de *phishing*" é um comportamento e,

neste estudo, trata-se do nosso objetivo de aprendizagem, que consiste em uma descrição de comportamentos que precisam ser aprendidos a partir de condições de ensino, tal como um jogo educativo. Sabemos que o ensino é eficiente quando constatamos que, antes dele, o aprendiz não conseguia emitir certo comportamento (objetivo de aprendizagem) e, em razão dele, o adquiriu ou aperfeiçoou, passando a lidar de modo mais efetivo com a sua realidade [Kienen et al., 2021].

Para aprendermos, é preciso um processo reiterado em que a ação é emitida sob certas condições e seguida por reforçamento [Moreira and Medeiros, 2018]. Contudo, reforçadores dificilmente são igualmente efetivos para todos e seu efeito depende de quão próximos temporalmente estão da ação sobre a qual devem repercutir [Moreira and Medeiros, 2018]. Nesse cenário, jogos educativos digitais são especialmente úteis porque promovem engajamento, expõem o aprendiz a situações-problema que requerem a apresentação dos comportamentos que devem ser aprendidos e, ainda, administram de modo imediato e automático estímulos, como pontos e *badges*, que, embora artificiais, tendem a funcionar como reforçadores [Panosso et al., 2015; Tsutsumi et al., 2020].

2.3. Jogo educativo

Jogos são sistemas artificiais delimitadas por regras em que o jogador se engaja voluntariamente, e que têm indicadores inequívocos de vitória e derrota. Tais sistemas se caracterizam pela interatividade, imersão e *feedback* automático [Gris and Souza, 2017; Gris et al., 2018]. Jogos educativos possuem todas as características dos "jogos" e, ainda, visam promover o aprendizado de comportamentos. Estudos sugerem que jogos educativos são eficientes para promover o aprendizado, sendo úteis para a capacitação de pessoas em larga escala [Panosso et al., 2015; Tsutsumi et al., 2020].

Baseados em elementos típicos de jogos e na teoria analítico-comportamental [Stutz, 2020; Gris et al., 2018; Tsutsumi et al., 2020], adaptamos *framework* típico de jogos [Schell, 2015] para orientar a criação do nosso jogo educativo (ver Anexo 1), que é composto por regras/mecânica (que explicitam objetivos de aprendizagem), narrativa, estética e tecnologia. Objetivos de aprendizagem são desenvolvidos na relação entre mecânica e regras. A mecânica consiste nas ações que o jogador deve emitir e as regras referem-se a quando e como essas ações devem ocorrer, para que aproximem o jogador da vitória. Essa relação, dentro do jogo, entre situação na qual é apropriado agir, ação que deve ser apresentada e resultado, é o próprio comportamento/objetivo que buscamos ensinar. A narrativa, por sua vez, é o modo de contar a história do jogo, tendo a finalidade de promover engajamento. A estética desempenha papel similar ao favorecer interesse e imersão. A tecnologia, por fim, é o meio que viabiliza a execução e interação com o jogo.

Adotamos neste estudo o *design* iterativo como procedimento para desenvolver o jogo, pois ele permite avaliações e melhorias contínuas através do teste sucessivo de protótipos funcionais [Gris and Souza, 2017]. A cada etapa concluída de desenvolvimento, segue-se para o teste do jogo, cuja finalidade é identificar o que precisa ser aperfeiçoado. Ao final desse processo, que não possui número exato de iterações, o jogo educativo deve se tornar eficiente na promoção de aprendizagem e de engajamento.

3. Trabalhos Relacionados

Revisamos a literatura sobre ensino de cibersegurança para mitigar ataques de *phishing* e descobrimos com Kovačević et al. (2020) que, mesmo estudantes da área de informática,

revelam dúvidas em relação a como se proteger de ciberataques. Contudo, nem por isso buscam aprender mais sobre cibersegurança, mostrando que todos precisam de algum incentivo para aprender e implementar boas práticas de segurança. Nessa linha, Rahman et al. (2020) identificaram em sua revisão da literatura que faltam capacitações em cibersegurança para crianças e que, antes disso, seus professores precisam ser treinados.

Com relação ao que ensinar sobre cibersegurança, Butavicius et al. (2015) mostraram que as pessoas são sensíveis a ataques que usam figuras de autoridade para conferir credibilidade a um e-mail com *phishing*, sendo importante treiná-las para essa situação. Por isso, em nosso jogo, investimos em mensagens de ameaça, que têm implícita a noção de autoridade. O estudo de Wash (2020) também fornece orientações sobre o que ensinar. Ele estudou os comportamentos de especialistas em TI para identificar e-mails com *phishing* e concluiu que é preciso desenvolver nas pessoas a capacidade de reconhecimento de sinais desse tipo de ataque. Por isso esse foi o nosso foco.

No que diz respeito a como ensinar, encontramos estudos promissores que usam jogos educativos. Farias et al. (2019), por exemplo, ensinaram conceitos de cibersegurança para crianças e adolescentes com o jogo "Self Protect" e obtiveram resultados positivos (97% de percepção de aprendizado). Arachchilage e Cole (2011), por sua vez, usaram um jogo desenvolvido com o Google App Inventor Emulator para ensinar como identificar e evitar URLs de e-mails com *phishing*. Essa pesquisa evidenciou melhora no comportamento de identificação de ameaças pelos usuários, mostrando, mais uma vez, que jogos podem ser úteis na educação em cibersegurança.

Com base nesses estudos, verificamos que ensinar cibersegurança é necessário [Rajasekharaiah et al., 2020] e que recursos tecnológicos podem viabilizar aprendizagens em larga escala. Por isso, decidimos desenvolver e avaliar a eficiência de jogo educativo para ensinar o comportamento de "distinguir e-mails legítimos de tentativas de *phishing*", respondendo a perguntas sobre aprendizado, autoconfiança, satisfação e usabilidade.

4. Método

4.1. Solução Computacional

Com base em nosso *framework* para jogos educativos, desenvolvemos ao longo de duas iterações as duas primeiras versões do jogo educativo Alerta. Priorizamos a criação da mecânica e das regras, bem como tornar o jogo funcional, cuidando da tecnologia. Nessas duas iterações, criamos uma narrativa simples para contar a história do Alerta.

Em termos de história, temos como contexto a empresa "JR Company", que está realizando processo seletivo para a vaga de assistente de segurança. O jogador participará dessa seleção, tendo como missão diferenciar e-mails legítimos de *phishing* recebidos por funcionários dessa empresa. O jogo possui 1 Pré-teste, 1 pós-teste e 4 fases de treino (para desenvolver os 4 objetivos de aprendizagem), cada uma com 3 tarefas de diferenciação de e-mails. Pré e pós possuem 12 tarefas de diferenciação cada um. O pré-teste é apresentado ao jogador como uma forma de a empresa identificar o que ele conhece sobre *phishing*. As 4 fases de treino são apresentadas como um treinamento para que os candidatos ao cargo tenham iguais oportunidades de sucesso na seleção. O pós-teste é apresentado como o teste que decidirá quem será contratado. O jogador precisa, então, obter mais de 100 pontos para ser aprovado, o que consiste na condição de vitória no jogo. A narrativa é realizada por meio de textos. Ao todo, criamos 36 e-mails, 18 legítimos e

18 *phishing*, sendo 6 de cada tipo para pré, pós e treino, os quais estão disponíveis neste *link* < http://dx.doi.org/10.13140/RG.2.2.26018.85448 >.

São apresentados no jogo 4 personagens: Hannah, Lucas, Joana e Mateus. Mateus, chefe do RH, é o primeiro que o jogador conhece, pois fornece as instruções do jogo. Os demais aparecem nas fases de treino, em que o jogador avalia suas informações (função na empresa e serviços que usa: banco, redes sociais, streaming e site de compra *online*) para, então, decidir se o e-mail recebido na caixa de entrada do personagem é legítimo ou tentativa de *phishing*. A ordem de progressão do jogo é a mesma para todos os jogadores.

Com relação às regras e mecânica do jogo, propusemos quatro objetivos de aprendizagem, constituintes do objetivo mais geral que é "Distinguir e-mails legítimos de tentativas de phishing", a saber: (1) Identificar se o e-mail é de empresa da qual o usuário é cliente ou de pessoa conhecida; (2) Identificar se o domínio do endereço de e-mail é compatível com aquele tipicamente usado pelo remetente; (3) Identificar se foi solicitada informação pessoal, clique em *link* ou *download* de arquivo; e (4) Identificar conteúdo de mensagem na qual o usuário é incentivado a tomar uma decisão rápida sob pena de sofrer punição. Para a emissão desses comportamentos, criamos a mecânica de clicar em botões para indicar se o e-mail apresentado na tela é legítimo ou contém *phishing*. No treino, se a marcação estiver correta, o jogador recebe a mensagem de que acertou; caso contrário, recebe uma mensagem de que errou. Esse *feedback* é imediato, tornando-o mais efetivo como reforçador. Cada fase de treino possui 3 e-mails e cada um pode ser respondido até duas vezes. O jogador tem, no máximo, seis tentativas para acertar as 3 tarefas. A tentativa consiste em uma resposta, legítimo ou *phishing*, para um e-mail. Ele precisa acertar, no mínimo, dois e-mails em cada fase, para que possa progredir. Do contrário, ele repete a fase quantas vezes for preciso. No caso do pré e do pós-teste, existe apenas a possibilidade de responder, sem que seja fornecido *feedback* de acerto ou erro. Além disso, nessas etapas do jogo, o jogador só pode avaliar cada e-mail uma vez.

Este jogo foi desenvolvido em C# na *engine* Unity 3D, dada a sua versatilidade e facilidade de uso [Unity Documentation, 2023]. Isso nos permitiu criar um jogo interativo e de fácil manutenção. A configuração de fases e a gestão dos dados dos e-mails foram implementadas com *Scriptable Objects*. A classe *PhaseConfig* foi utilizada para armazenar os dados dos e-mails, incluindo remetente, conteúdo e um indicador booleano que identifica se o e-mail é *phishing*. Implementamos um campo para definir a pontuação necessária para avançar para a próxima fase. A classe *EmailManager* foi responsável por gerenciar a exibição dos e-mails, verificar as respostas do jogador e atualizar a pontuação. Os resultados de cada teste e fase de treino são salvos em um arquivo de texto local. A interface do jogo foi projetada para ser intuitiva. E-mails não lidos têm uma coloração vermelha, enquanto os lidos ficam verdes. Mensagens de *feedback* aparecem em *pop-ups* e só podem ser fechadas mediante clique do usuário no botão correspondente. Exemplos de telas do jogo podem ser consultadas no Anexo 3. O código desenvolvido está disponível no GitHub < https://github.com/JassonJr1/alerta_prototype >.

4.2. Avaliação Experimental

4.2.1. Participantes

Pesquisa piloto - 1ª Iteração. Para melhorar o método de coleta de dados, foi realizada pesquisa piloto com quatro alunos veteranos do curso de computação (P1 a P4). **Pesquisa principal - 2ª Iteração.** Participaram 27 pessoas, P5 a P27, com média de idade de 25,81

anos (*DP* = 7,50), variando de 20 a 49, sendo 63% homens. A escolaridade predominante foi ensino superior em andamento (66,70%), e os participantes, tipicamente, eram alunos de Computação (37,00%), Direito (22,21%) ou Odontologia (11,10%). Para a análise, consideramos apenas os dados de participantes cuja nota no pré-teste foi menor que 85%, para atenuar o efeito de teto. Excluímos P5, P9, P12, P15, P17, P18, P22, P25 e P30, cuja pontuação média no Pré-teste foi de 94,44. A amostra final ficou com 18 participantes. Todos assinaram Termo de Consentimento Livre e Esclarecido (TCLE). Este estudo foi aprovado por Comitê de Ética em Pesquisa, CAAE n. 77411823.6.0000.5302.

4.2.2. Instrumentos

Utilizamos os seguintes instrumentos: (1) Caracterização da amostra: avaliava idade, gênero, escolaridade e nome do curso, no caso de universitários ou graduados; (2) Medida de aprendizagem: composta por 12 tarefas de avaliação de e-mails, no pré e no pós, todas compondo o jogo Alerta. Criamos também uma escala para mensurar o grau de segurança do jogador em suas respostas de classificação de e-mails, a qual era preenchida em uma folha e variava de "1 - Totalmente Inseguro" a "4 - Totalmente Seguro"; (3) Avaliação de Usabilidade (PSSUQ, ver Anexo 2): trata-se de ferramenta adotada na literatura [Vlachogianni et al., 2023], tendo 15 itens que avaliam utilidade do sistema (Itens 1 a 6), qualidade da informação (Itens 7 a 12) e qualidade da interface (Itens 13 a 15). As respostas são dadas em uma escala Likert de concordância, "1 - Discordo Totalmente" a "5 - Concordo Totalmente". Realizamos a tradução desse instrumento, mas sem a condução de um estudo de adaptação transcultural; (4) Teste de Satisfação: composto por 5 itens que avaliavam o jogo, também em uma escala Likert de concordância.

4.2.3 Procedimento de coleta e análise de dados

Na pesquisa piloto, correspondente à 1ª iteração de desenvolvimento, pedimos que todos assinassem o TCLE e fornecemos instruções sobre a história e o objetivo do jogo. A partir dos *feedbacks* recebidos dos quatro participantes e das observações que fizemos, implementamos melhorias no jogo, de modo a construir a segunda versão. Alteramos a forma como a história é contada, integrando-a a um diálogo com o chefe do RH. Adicionamos uma sinalização visual para indicar quais e-mails foram ou não lidos e definimos a regra de, pelo menos, dois acertos no treino para poder progredir de fase. Mudamos a apresentação do início das fases, incluindo uma identificação da fase em que o jogador se encontra e informativos para lembrá-los sobre a função de cada botão.

Nessa 1ª iteração obtivemos evidências da eficiência do jogo. Nenhum participante precisou das 6 tentativas para distinguir corretamente os e-mails das fases de treino, e o número de tentativas utilizadas caiu ao longo das fases. Observamos, ainda, que, na Fase 1, a média de acertos foi de 83,34, na Fase 2, foi 100, na Fase 3, foi 91,67 e, na Fase 4, foi 100. Tais resultados sugerem que os objetivos de aprendizagem foram corretamente apresentes durante o treino em função dos *feedbacks* de acerto e erro programados. Quando comparamos pré e pós, notamos que os 4 participantes melhoram de desempenho. A média no pré foi de 68,75% de acertos e de 100% no pós. Esses dados evidenciam a eficiência do jogo na promoção do objetivo de aprendizagem. Finalmente, os escores médios de usabilidade (*Média* = 4,71; *DP* = 0,11) e de satisfação (*Média* = 4,75; *DP* = 0,19) foram elevados, com pequenos desvios padrão.

Para a coleta principal, 2ª iteração no desenvolvimento do jogo, pedimos que todos lessem e assinassem o TCLE. Em seguida, foi explicado o objetivo do jogo Alerta e sua história. O pesquisador lembrou que não forneceria respostas, mas poderia ajudar

em caso de falha no funcionamento do jogo. Foi solicitada atenção à tarefa e proibida a interação com outros participantes. Foi explicado, ainda, que, para cada e-mail examinado no pré ou pós-teste, era preciso preencher uma folha indicando o grau de segurança na própria resposta. Os participantes foram incentivados a vencer o jogo e avisados de que ao final seriam solicitados a responder uma avaliação sobre a sua experiência. Após as instruções, todos começaram a jogar o Alerta em um ambiente controlado de laboratório. A pesquisa foi conduzida coletivamente, tendo duração média de 1 hora. Nesse tempo, os participantes jogaram pré-teste, 4 fases do jogo e pós-teste. Os dados coletados foram analisados de modo qualiquantitativo, enfatizando estatísticas descritivas e o teste dos postos sinalizados de Wilcoxon para comparar pré e pós. A íntegra dos dados está disponível neste *link* < http://dx.doi.org/10.13140/RG.2.2.16162.24004 >.

5. Resultados e Discussão

A Tabela 1 exibe os resultados das quatro fases de treino. Notamos melhora sutil de desempenho dos participantes à medida que avançavam nessas fases, partindo de uma média de 90,74% de acertos na Fase 1 para 98,15 na Fase 4. Cada vez menos erros ocorreram e menos tentativas foram necessárias para progredir nas fases. Alguns participantes (P11, P16, P19 e P21), que cometeram erros, não usaram todas as tentativas à disposição, mas conseguiram avançar de fase porque haviam acertado o mínimo de dois e-mails. Destacamos que o computador do P28 desligou quando estava iniciando a repetição da 1ª fase, após não ter atingido o critério mínimo. Apenas os dados do pré-teste ficaram armazenados. Pela programação do jogo, ele precisou repetir o pré-teste, mas não analisamos esses dados. Depois, repetiu a Fase 1, que já seria necessária.

Tabela 1. Resultados dos participantes nas 4 fases da 2ª iteração do jogo.

Participante	Fase 1		Fase 2		Fase 3		Fase 4	
	Acertos	Tentativas	Acertos	Tentativas	Acerto	Tentativas	Acerto	Tentativas
P6	100,00	3	100,00	5	100,00	4	100,00	4
P7	100,00	3	100,00	3	100,00	3	100,00	3
P8	100,00	3	100,00	3	100,00	3	100,00	3
P10	66,67	4	100,00	3	100,00	4	100,00	3
P11	66,67	3	66,67	3	66,67	3	66,67	3
P13	100,00	5	100,00	4	100,00	4	100,00	3
P14	100,00	4	100,00	3	100,00	3	100,00	3
P16	66,67	3	66,67	3	66,67	3	100,00	3
P19	66,67	3	66,67	3	66,67	3	100,00	3
P20	100,00	5	100,00	3	100,00	3	100,00	3
P21	100,00	3	66,67	3	100,00	4	100,00	3
P23	100,00	6	100,00	6	100,00	3	100,00	4
P24	100,00	4	100,00	3	100,00	4	100,00	3
P26	100,00	4	100,00	3	100,00	4	100,00	3
P27	100,00	4	100,00	3	100,00	3	100,00	3
P28	100,00	6	100,00	4	100,00	4	100,00	3
P29	100,00	3	100,00	3	100,00	3	100,00	3
P31	66,67	5	66,67	4	100,00	4	100,00	3
Média	**90,74**	**3,94**	**90,74**	**3,44**	**94,45**	**3,44**	**98,15**	**3,11**
DP	**15,36**	**1,06**	**15,36**	**0,86**	**12,78**	**0,51**	**7,86**	**0,32**

Esses dados mostraram que os 4 objetivos de aprendizagem foram consistentemente emitidos e reforçados. Tal como no jogo de Farias et al. (2019), obtivemos evidências favoráveis de que o treinamento do Alerta foi adequado para que pessoas com diferentes repertórios pudessem realizá-lo, progredindo sem dificuldades pelas fases do jogo. O Alerta avançou em relação a outros jogos porque explicitou objetivos de aprendizagem e construiu o jogo em função deles [Panosso et al., 2015]. Com os dados disponíveis, ainda não conseguimos explicitar por qual motivo um participante precisou de mais tentativas em uma fase ou em relação aos demais. Hipóteses a serem investigadas em estudos futuros são: (a) cansaço e/ou falta de atenção; e (b) complexidade do comportamento sendo treinado.

A Tabela 2 exibe os resultados no pré e no pós-teste. O desempenho médio do pré-teste foi de 73,61 (*DP* = 15,46) e aumentou para 89,35 (*DP* = 8,95), sendo este achado estatisticamente significativo e com tamanho do efeito alto ($W = 153{,}00$; $z = 3{,}621$; $p < 0{,}001$; $r = 1{,}00$; $EP = 0{,}269$). Esse resultado decorre do treinamento bem-sucedido nas 4 fases e consiste em uma evidência da eficiência do jogo no ensino do comportamento de "diferenciar e-mails legítimos de tentativas de *phishing*". Podemos notar pelos dados do pré-teste que todos os participantes já tinham parte desse comportamento em seu repertório e P19 não exibiu melhora (talvez por não ter treinado o suficiente cf. Tabela 1). Não obstante, no geral, o jogo foi útil para promover aperfeiçoamento desse repertório, que é uma forma de aprendizado. Portanto, tecnologias como esta podem aperfeiçoar comportamentos de cibersegurança e, possivelmente, mitigar ataques de *phishing*, tal como demandado pela literatura [Rahman et al., 2020; Guilherme et al., 2021].

Tabela 2. Percentual de acertos no pré e pós-teste da 2ª iteração do jogo.

Participante	% Acertos		Diferença Pré-Pós	Participante	% Acertos		Diferença Pré-Pós
	Pré	Pós			Pré	Pós	
P6	50,00	75,00	25,00	P20	25,00	75,00	50,00
P7	83,33	91,67	8,33	P21	75,00	91,67	16,67
P8	83,33	100,00	16,67	P23	83,33	91,67	8,33
P10	75,00	100,00	25,00	P24	75,00	83,33	8,33
P11	58,33	75,00	16,67	P26	75,00	83,33	8,33
P13	83,33	91,67	8,33	P27	83,33	91,67	8,33
P14	66,67	100,00	33,33	P28	75,00	83,33	8,33
P16	83,33	91,67	8,33	P29	83,33	100,00	16,67
P19	83,33	83,33	0,00	P31	83,33	100,00	16,67
Média Pré-teste (DP)			**Média Pós-teste (DP)**				**Média de Diferença (DP)**
73,61 (*DP* = 15,46)			89,35 (*DP* = 8,95)				15,74 (*DP* = 11,75)

A Tabela 3 exibe o registro do grau de segurança em relação às respostas fornecidas no pré e pós-teste, sendo que o dado de P16 não está disponível porque ele preencheu de modo incorreto esse instrumento, tendo avaliado a segurança que o e-mail passava para ele e não a segurança que sentia nas próprias respostas. Podemos notar no grupo uma melhora de 3,22 (*DP* = 0,70) de média de segurança no pré-teste para 3,60 (*DP* = 0,45) no pós-teste, com redução do desvio padrão, sugerindo maior homogeneidade nos escores de segurança. Essa mudança, muito próxima da pontuação máxima (4,00), foi estatisticamente significativa e apresentou elevado tamanho do efeito ($W = 89{,}00$; $z = 2{,}291$; $p = 0{,}024$; $r = 0{,}695$; $EP = 0{,}294$), mais uma vez indicando a eficiência do Alerta, agora em relação à promoção de autoconfiança.

Tabela 3. Grau de segurança nas respostas ao pré e pós-teste da 2ª iteração do jogo.

Participante	% Acertos		Diferença Pré-Pós	Participante	% Acertos		Diferença Pré-Pós
	Pré	Pós			Pré	Pós	
P6	2,58	3,67	1,08	P20	1,00	3,00	2,00
P7	3,25	4,00	0,75	P21	3,33	3,42	0,08
P8	3,25	4,00	0,75	P23	3,75	3,75	0,00
P10	3,50	3,67	0,17	P24	3,83	3,92	0,08
P11	2,75	3,18	0,43	P26	3,67	3,58	-0,08
P13	2,75	3,83	1,08	P27	3,42	3,42	0,00
P14	3,50	4,00	0,50	P28	3,00	2,33	-0,67
P16	---	---	---	P29	4,00	4,00	0,00
P19	3,50	4,00	0,50	P31	3,67	3,42	-0,25
Média Pré-teste (DP)			**Média Pós-teste (DP)**				**Média de Diferença (DP)**
3,22 (*DP* = 0,70)			3,60 (*DP* = 0,45)				0,38 (*DP* = 0,63)

Esses dados sugerem que o jogo Alerta pode ser útil para reduzir dúvidas de usuários sobre *phishing* e favorecer, inclusive, a sua percepção de ser capaz de lidar de modo autônomo com a tarefa de identificar e-mails com *phishing*, algo que, por vezes, mesmos estudantes de informática não sentem [Kovačević et al., 2020]. Lembramos que P23, P26, P28 e P31 não exibiram melhora, o que aponta para a necessidade de aperfeiçoamento do jogo e dos dados coletados, para que possamos entender esses casos. Em síntese, os dados de aperfeiçoamento do comportamento de "distinguir e-mails legítimos de tentativas de *phishing*" e de autoconfiança sugerem que o jogo educativo Alerta foi eficiente no desenvolvimento de elementos básicos do senso crítico, proposto por Wash (2020) como crucial para que a pessoa saiba evitar ataques de *phishing*.

Tabela 4. Resultados de usabilidade e satisfação da 2ª iteração do jogo.

Par.	Escores				Par.	Escores			
	Utilidade	Informação	Interface	Satisfação		Utilidade	Informação	Interface	Satisfação
P6	4,83	4,83	5,00	5,00	P20	5,00	4,50	5,00	5,00
P7	5,00	5,00	5,00	5,00	P21	5,00	4,83	5,00	5,00
P8	5,00	5,00	5,00	5,00	P23	4,83	4,67	5,00	4,80
P10	5,00	4,00	5,00	5,00	P24	5,00	5,00	4,67	5,00
P11	5,00	4,17	5,00	5,00	P26	4,83	4,83	5,00	5,00
P13	3,00	2,67	3,67	4,00	P27	4,67	4,00	4,67	4,80
P14	4,33	2,83	3,67	4,60	P28	3,00	3,00	3,00	3,00
P16	4,33	4,67	3,67	4,40	P29	5,00	5,00	5,00	5,00
P19	4,50	4,33	5,00	5,00	P31	5,00	5,00	5,00	5,00
Média Usabilidade (DP): 4,52 (*DP* = 0,65)					**Média Satisfação (DP):** 4,76 (*DP* = 0,52)				

Por fim, a Tabela 4 exibe os dados relativos aos instrumentos de avaliação de percepção de usabilidade e de satisfação, ambos com pontuação máxima igual a 5. Podemos notar que tanto a média de usabilidade (4,52), quanto a de satisfação (4,76) foram elevadas, sugerindo que o jogo educativo foi fácil de ser utilizado e gerou sentimentos e percepções positivas, o que é uma medida indireta do engajamento que pode promover. Especificamente, no caso da usabilidade, podemos notar que nas suas três dimensões avaliativas os resultados também foram positivos: utilidade apresentou média de 4,63 (*DP* = 0,64), informação foi 4,35 (*DP* = 0,78) e interface 4,63 (*DP* = 0,64).

Esses dados sugerem, portanto, que o jogo Alerta, além de eficiente na promoção de aprendizagens e autoconfiança, tem potencial para gerar engajamento, cumprindo, portanto, os dois requisitos básicos de um jogo educativo [Tsutsuni et al., 2020].

Em conjunto, os dados de aprendizagem, grau de segurança, usabilidade e satisfação sugerem que o jogo Alerta é promissor. Não obstante, existem limitações em nosso estudo que devem ser consideradas: (1) pequeno número de participantes e com baixa variabilidade de características e repertórios de cibersegurança, além de, em geral, terem obtido elevado percentual de acertos no pré-teste. Para estudos futuros, sugerimos coletar dados com participantes mais diversos, que apresentem menos de 60% de acertos no pré-teste, para que possamos ter um análise mais robusta da eficiência do jogo; (2) pequeno escopo dos objetivos de aprendizagem abarcados pelo jogo, sendo importante incluir objetivos mais complexos e que capacitem as pessoas para lidar com técnicas de *phishing* mais desafiadoras, como *spear phishing*; (3) uso de instrumentos sem evidências psicométricas robustas, o que requer estudos específicos de construção e adaptação de medidas para pesquisa e intervenção na subárea de educação em cibersegurança.

Ademais, sugerimos que a terceira iteração de desenvolvimento do jogo contemple os seguintes aperfeiçoamentos: (1) obrigue o jogador a usar todas as tentativas que possui para tentar acertar as tarefas do treino de modo que o comportamento esperado possa ser emitido e reforçado, aspecto necessário para o aprendizado; (2) implemente a medida de autoconfiança no jogo, como parte do pré e pós-teste; (3) implemente aperfeiçoamentos na estética e narrativa, para tornar o jogo mais interessante e, assim, atrair jogadores e fazê-los desejarem jogá-lo; (5) aperfeiçoar a usabilidade para contemplar pessoas com dificuldades em leitura; (6) aprimorar os *feedbacks* fornecidos durante o jogo, para aumentar o seu valor reforçador. Com esses aprimoramentos, esperamos que o Alerta se torne uma ferramenta educativa ainda mais eficiente.

6. Conclusão

O objetivo deste estudo foi avaliar a eficiência do jogo educativo Alerta para o ensino do comportamento de "distinguir e-mails legítimos de tentativas de *phishing*". A partir da criação de um protótipo de jogo, ao longo de duas iterações, que priorizaram desenvolvimento de regras, mecânica e tecnologia, identificamos resultados promissores. Os participantes conseguiram progredir pelas fases de treino sem dificuldades e com melhoras sutis de desempenho. Na comparação entre pré e pós-teste, encontramos evidências de aprendizado dos objetivos propostos para o jogo, bem como de desenvolvimento do sentimento de autoconfiança em relação à capacidade de reconhecer tentativas de *phishing*. Finalmente, notamos que o jogo foi percebido como fácil de usar, além de prover uma experiência satisfatória. Tais resultados vêm ao encontro da nossa proposta inicial, que era criar um sistema por meio do qual as pessoas pudessem aprender comportamentos relevantes de modo confortável e prazeroso, sendo simples para elas, mesmo sem auxílio externo, fazer o *download* do jogo e jogá-lo.

Sugerimos que estudos futuros avancem em relação a esta pesquisa, buscando lidar com as limitações que apresentamos. Recomenda-se, especialmente, coletar dados com um grupo mais diversificado de participantes e ampliar o escopo dos objetivos de aprendizagem desenvolvidos por meio do jogo, para que ele possa ajudar a sociedade a mitigar ataques de *phishing*, cada vez mais frequentes e elaborados. Consideramos também promissor testar o jogo com populações específicas, a exemplo dos idosos.

7. Referências

Alanazi, M., Freeman, M., and Tootell, H. (2022). Exploring the factors that influence the cybersecurity behaviors of young adults. *Computers in Human Behavior*, *136*(107376), 1-14. https://doi.org/10.1016/j.chb.2022.107376

APWG. (2023). Phishing activity trends reports. Recuperado de https://bit.ly/4cEA85K

APWG. (2024). Phishing activity trends reports. Recuperado de https://bit.ly/4dC6GgW

Butavicius, M., Parsons, K., Pattinson, M., and McCormac, A. (2016). Breaching the human firewall: Social engineering in phishing and spear-phishing emails. https://doi.org/10.48550/ARXIV.1606.00887

Centro de Estudos, Resposta e Tratamento de Incidentes de Segurança no Brasil. (2024). *Páginas Falsas Utilizadas em Tentativas de Phishing*. Recuperado de https://stats.cert.br/phishing/

Diorio, R. F., Serafim, E., Alves, K. R., and Meira, M. C. (2018). Segurança da Informação e de Sistemas Computacionais: Um Estudo Prático sobre Ataques Utilizando Malwares. Anais SULCOMP, 9. Recuperado de https://bit.ly/45Oj62J

Documentation. [s.d.]. Unity documentation. Recuperado de https://docs.unity.com/

Farias, F. L. de O., Medeiros, N. A. A. de, Rocha, S. L. da, Medeiros, D. F. de, Nóbrega, E. C. da, Burlamaqui, A., and Madeira, C. (2019). Self Protect: Um jogo para auxílio no ensino de conceitos relacionados a Segurança na Internet para Crianças e Adolescentes. In *Anais do XXV Workshop de Informática na Escola (pp 246-255). Sociedade Brasileira de Computação.* https://doi.org/10.5753/cbie.wie.2019.246

Gris, G., and De Souza, S. R. (2016). Digital educational games and model of network relations: Development and evaluating of the physical prototype of Korsan game. *Perspectivas em Análise do Comportamento*, *7*(1), 114–132. https://doi.org/10.18761/pac.2016.003

Guilherme, L. P., Ferreira, M. F., Da Fonseca, G. M., and Lazarin, N. M. (2021). Uma breve noção sobre o comportamento dos internautas em relação à segurança na rede. In *Escola Regional de Sistemas de Informação do Rio de Janeiro (ERSI-RJ) (pp. 1-7)*. Sociedade Brasileira de Computação. https://doi.org/10.5753/ersirj.2021.16972

Hartwig, K., and Reuter, C. (2021). Nudge or restraint: How do people assess nudging in cybersecurity - A representative study in Germany. In *Proceedings of the 2021 European Symposium on Usable Security, (pp 141-150)*. ACM. https://doi.org/10.1145/3481357.3481514

Kienen, N., Panosso, M. G., Nery, A. G. S., Waku, I., and Carmo, J. S. (2021). Contextualização sobre a programação de condições para desenvolvimento de comportamentos (PCDC): Uma experiência brasileira. *Perspectivas em Análise do Comportamento*, *12*(2), 360-390. https://doi.org/10.18761/PAC.2021.jul110

Kovacevic, A., Putnik, N., and Toskovic, O. (2020). Factors related to cyber security behavior. *IEEE Access*, *8*, 125140–125148. https://doi.org/10.1109/access.2020.3007867

Melo, L. P., Amaral, D. M., Sakakibara, F., De Almeida, A. R., De Sousa Jr, R. T., and Nascimento, A. (2011). Análise de Malware: Investigação de Códigos Maliciosos

Através de uma Abordagem Prática. In *Minicursos do XI Simpósio Brasileiro de Segurança da Informação e de Sistemas Computacionais (pp. 9–52)*. Sociedade Brasileira de Computação. https://doi.org/10.5753/sbc.9559.1.1

Moreira, B. M., and Medeiros, A. C. (2018). *Princípios básicos de análise do comportamento*. Porto Alegre: Artmed.

Morgado, E. M., Távora, C. G., Lima, A. C. P. F. de, Albino, J. P., and Bucchianico, I. (2023). Caso de Cyber Fraud por telefone no Brasil e A Inteligência Artificial: Vítimas Idosas, Spoofing até a Manipulação por Engenharia Social. *In Inteligência Artificial e suas Aplicações Interdisciplinares (pp. 113–126)*. Editora e-Publicar. https://doi.org/10.47402/ed.ep.c202321007201

Panosso, M. G., Souza, S. R., and Haydu, V. B. (2015). Características atribuídas a jogos educativos: Uma interpretação Analítico-Comportamental. *Revista Quadrimestral da Associação Brasileira de Psicologia Escolar e Educacional*, *19*(2), 233-241. http://dx.doi.org/10.1590/2175-3539/2015/0192821

Rahman, N. A. A., Sairi, I. H., and Zizi, N. A. M., Khalid, F. (2020). The importance of cybersecurity education in school. *International Journal of Information and Education Technology*, *10*(5), 378–382. https://doi.org/10.18178/ijiet.2020.10.5.1393

Rajasekharaiah, K. M., Dule, C. S., and Sudarshan, E. (2020). Cyber security challenges and its emerging trends on latest technologies. *IOP Conference Series: Materials Science and Engineering*, *981*(2), 022062. https://doi.org/10.1088/1757-899x/981/2/022062

Ruoslahti, H., Coburn, J., Trent, A., and Tikanmäki, I. (2021). Cyber skills gaps – A systematic review of the academic literature. *Connections The Quarterly Journal*, *20*(2), 33–45. https://doi.org/10.11610/connections.20.2.04

Scarfone, K. A., Souppaya, M. P., Cody, A., and Orebaugh, A. D. (2008). *Technical guide to information security testing and assessment*. Gaithersburg, MD: National Institute of Standards and Technology. https://doi.org/10.6028/nist.sp.800-115

Schell, J. (2015). *The art of game design:* A book of lenses. CRC Press.

Skinner, B. F. (1981). *Ciência e comportamento humano*. São Paulo: Martins Fontes.

Sophos. (2023). The state of ransomware 2023. Recuperado de https://bit.ly/3xROD79

Souza, L. C., and Tanaka, S. S. (2023). Estudo sobre ataques de phishing e suas técnicas de defesa. *Revista Terra & Cultura: Cadernos de Ensino e Pesquisa*, *39*(Especial), 90–95. Retrieved November 18, 2023, from https://bit.ly/4cOisV2

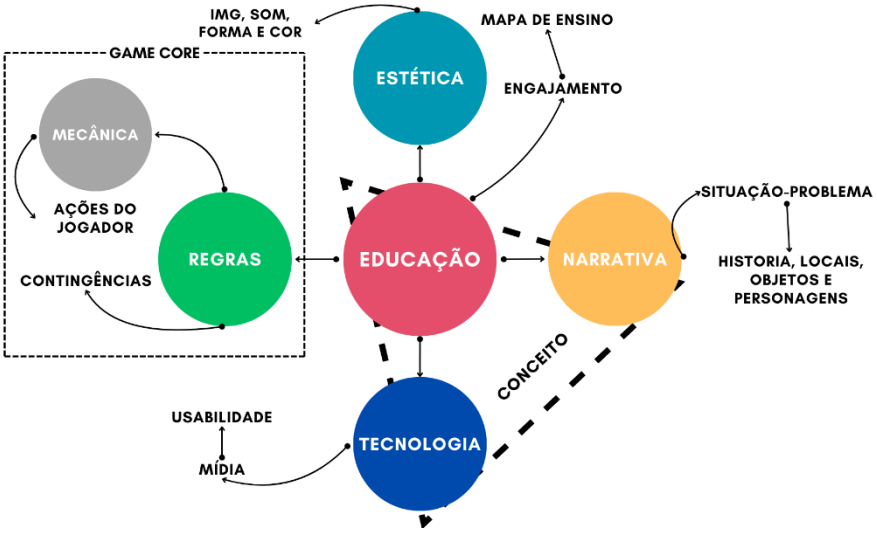
Syafitri, W., Shukur, Z., Mokhtar, U. A., Sulaiman, R., and Ibrahim, M. A. (2022). Social engineering attacks prevention: A systematic literature review. *IEEE Access*, *10*, 39325–39343. https://doi.org/10.1109/access.2022.3162594

Tsutsumi, M. M. A., Goulart, P. R. K., Silva Júnior, M. D., Haydu, V. B., and Jimenéz, É. L. de O. (2020). Avaliação de jogos educativos no ensino de conteúdos acadêmicos: Uma revisão sistemática da literatura. *Revista Portuguesa de Educação*, *33*(1), 38–55. https://doi.org/10.21814/rpe.19130

Wash, R. (2020). How experts detect phishing scam emails. *Proceedings of the ACM on Human-Computer Interaction*, *4*(CSCW2), 1–28. https://doi.org/10.1145/3415231

8. Anexos

Anexo 1 - Framework para desenvolvimento de jogos educativos.



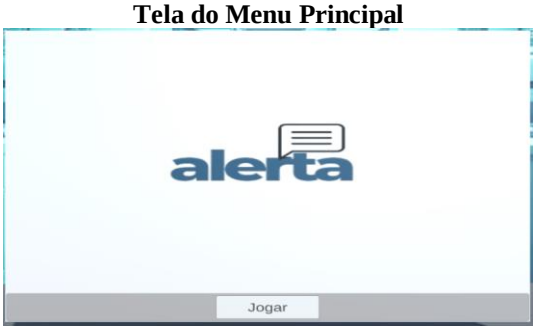
Anexo 2 - Avaliação de Percepção de usabilidade e de Satisfação.

Avaliação de Percepção de Usabilidade
1 . No geral, estou satisfeito com o quão fácil é usar este sistema.
2 . Foi simples usar este sistema.
3. Consegui concluir as tarefas rapidamente usando este sistema.
4. Eu me senti confortável usando este sistema.
5. Foi fácil aprender a usar este sistema.
6. Eu acredito que eu poderia me tornar produtivo rapidamente usando este sistema.
7. O sistema apresentou mensagens de erro que claramente indicaram como corrigir problemas.
8. Sempre que cometi algum erro, eu pude recuperar de forma fácil e rápida.
9. As informações (como tutoriais, mensagens na tela e outras documentações) fornecidas com este sistema eram claras.
10. Foi fácil encontrar as informações que eu precisava.
11. As informações foram eficazes para me ajudar a concluir as tarefas e cenários.
12. A organização das informações nas telas do sistema estava clara.
13. A interface do sistema é agradável.
14. Eu gostei de usar a interface deste sistema.
15. No geral, estou satisfeito com este sistema.

Avaliação de Satisfação
1. No geral, estou satisfeito com a história do jogo.
2. No geral, estou satisfeito com a narrativa do jogo.
3. No geral, estou satisfeito com as regras do jogo.
4. No geral, estou satisfeito com o modo de jogar.
5. No geral, estou satisfeito com o jogo Alerta.

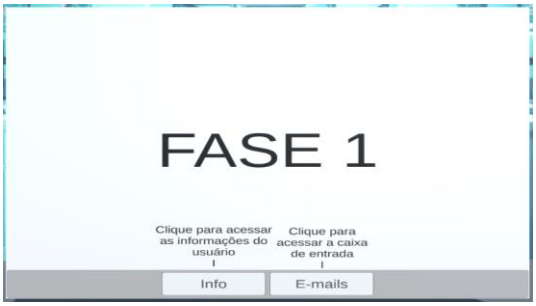
Anexo 3 - Principais telas do jogo Alerta.

Tela do Menu Principal

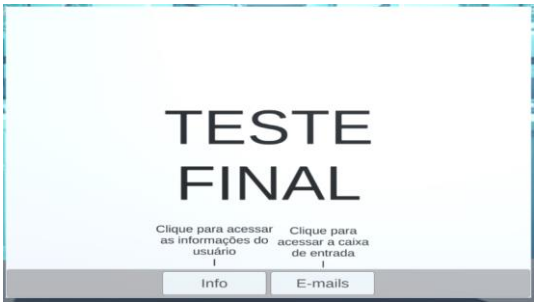


Tela de Instrução

Tela de Início Padrão das Fases

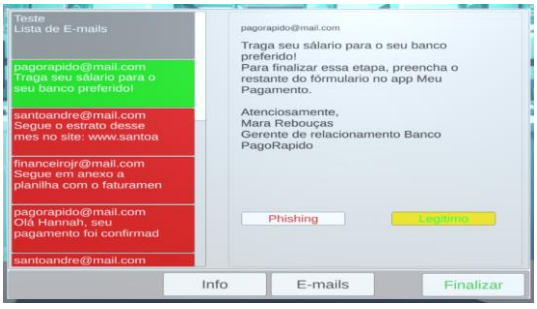


Tela de Início Padrão dos Testes

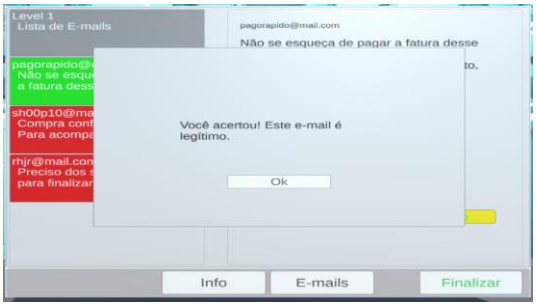


Tela de Info do Personagem

Tela de Caixa de Entrada e com E-mail



Tela de *Feedback* sobre acerto



Tela de *Feedback* sobre erro