

MATEUS KUHN CHEDID

OTIMIZAÇÃO DE CAMPANHAS DE MARKETING DIGITAL
UTILIZANDO DELINEAMENTO DE EXPERIMENTOS

Trabalho de Formatura apresentado à
Escola Politécnica da Universidade de
São Paulo para obtenção do Diploma
de Engenheiro de Produção.

São Paulo
2020

MATEUS KUHN CHEDID

OTIMIZAÇÃO DE CAMPANHAS DE MARKETING DIGITAL
UTILIZANDO DELINEAMENTO DE EXPERIMENTOS

Trabalho de Formatura apresentado à
Escola Politécnica da Universidade de
São Paulo para obtenção do Diploma
de Engenheiro de Produção.

Orientador: Prof. Dr. Davi Noboru
Nakano

São Paulo
2020

FICHA CATALOGÁFICA

Chedid, Mateus Kuhn

Otimização de Campanhas de Marketing Digital Utilizando Delineamento de Experimentos / M. K. Chedid -- São Paulo, 2020.

100 p.

Trabalho de Formatura - Escola Politécnica da Universidade de São Paulo. Departamento de Engenharia de Produção.

1.Delineamento Experimental 2.Experimento Fatorial Fracionado
3.Marketing Digital 4.Mobile Games I.Universidade de São Paulo. Escola Politécnica. Departamento de Engenharia de Produção II.t.

Ao maior politécnico que conheci:
Ricardo Laquis Chedid, meu avô.

AGRADECIMENTOS

Agradeço a todos aqueles que possibilitaram ou engrandeceram minha graduação.

À minha mãe, Kátia Andréia Kuhn Chedid, que me ensina a sempre querer e buscar aprender. Ao meu pai, Rogério de Sá Chedid, que me inspira a construir a vida com honestidade e paixão. Obrigado por sempre fazerem o que podiam e me darem mais do que tinham. Tudo que sou devo a vocês. Aos meus irmãos Rodrigo Kuhn Chedid e Leonardo Kuhn Chedid, pela cumplicidade silenciosa e pela segurança de nunca se estar sozinho. Ao meu avô e politécnico, Ricardo Laquis Chedid, que me mostrou que o amor pode mudar o imutável, que a humildade nos enobrece e que o perdão pode ser fácil. À minha família, pelas risadas em momentos tristes, pelas brigas em eventos felizes e por sempre garantirem que estejamos juntos e festejando.

A todos os amigos que trouxe comigo e que fiz pelo caminho. Aos amigos do colégio, por me mostrarem que, quando o carinho é verdadeiro, o tempo e a distância nunca são obstáculos. Aos amigos que fiz durante meu intercâmbio em Turim, por serem minha família na nova casa e, principalmente, por me segurarem quando eu podia cair. E aos amigos politécnicos, pelos melhores anos da minha vida.

Ao Centro Acadêmico da Engenharia de Produção (CAEP), por me acolher e ser minha casa e minha família na Escola Politécnica. Por ter sido o palco para minhas melhores recordações, a base para minhas maiores amizades e o motivo para meu amor pela Poli. A todas as pessoas maravilhosas que fizeram e fazem parte desse grupo. Que muitos ainda tenham a sorte que eu tive de passar pela graduação com tanto amor e companheirismo.

E, claro, à Adelaide Cristina Carvalho Costa e ao Osni Alves dos Santos, minha segunda mãe e meu segundo pai. Por sempre estarem presentes e dispostos a ajudar. Por acolherem a todos e fazerem de tudo. Por terem os maiores corações que eu conheço.

À Profa. Dra. Linda Lee Ho, pelo apoio e aconselhamento. E ao Prof. Dr. Davi Noboru Nakano, por me guiar e acompanhar nesse processo e, mais importante, pela paciência, compreensão e companheirismo.

Aos meus colegas na Wildlife, que, além de grandes profissionais, têm sido grandes amigos.

À saúde mental, infelizmente ainda tão escassa nessa Escola, e à terapia, por salvar a minha graduação (e a de muitos politécnicos). Que nossa instituição saiba evoluir com humanidade e consciência. À arte, à cultura e ao entretenimento que hoje, mais do que nunca, se mostram essenciais para vivermos saudáveis.

Às politécnicas e aos politécnicos LGBTQIA+. Às mulheres, aos pretos, aos pobres e a qualquer representante de minorias na Poli. Aos que vieram antes e aos que estiveram ao meu lado, muito obrigado. Aos que estão por vir, que não tenham medo, que ocupem os espaços e que sejam sempre mais. Por uma Engenharia mais diversa e inclusiva.

Ao Ensino Público, por existir, resistir e sempre mostrar seu valor. Que cresça e que seja feito para todas e todos.

Ain't about how fast I get there
Ain't about what's waiting on the other side
It's the climb
(Miley Cyrus)

RESUMO

O trabalho elaborado se propõe a criar uma solução real e prática para o processo de otimização de campanhas de marketing de uma empresa de *mobile games*. O processo atual trata as informações de forma binária e os fatores como independentes entre si. Por meio de um modelo de Experimento Fatorial Fracionado 2^{k-p} , comparou-se o modelo atual com o proposto a fim de definir-se qual é mais vantajoso para a empresa.

Neste trabalho são utilizados conceitos de Delineamento Experimental para se elaborar um modelo de testagem de vídeos de marketing com iterações de elementos distintos que possa ser colocada em prática pela empresa. Inicialmente, as características relevantes de um Experimento Fatorial Fracionado 2^{k-p} são detalhadas. O mercado de *mobile games*, o posicionamento da empresa nele e o processo atual e suas particularidades são, então, expostos. A seguir, são estruturados e modelados dois projetos equivalentes para análise de um mesmo vídeo com as mesmas alterações realizadas. O primeiro é exatamente como a empresa faz esse tipo de análise atualmente, e o segundo segue o modelo proposto. Os dois projetos foram executados e seus desempenhos no resultado final foram analisados e comparados.

Quanto aos resultados, o modelo proposto mostrou um desempenho superior, menos dependente de intuição e de conhecimentos prévios não comprovados. Apesar de seus custos de produção e testagem serem superiores, o modelo proposto apresentou resultados suficientes para compensar essa diferença. Utilizando o processo baseado no Delineamento Experimental, a empresa pode otimizar as performances de seus anúncios, o que reduz consideravelmente o custo de operação e aumenta o resultado geral da empresa.

Palavras Chave: Delineamento Experimental; Experimento Fatorial Fracionado; Marketing Digital; Jogos Digitais

ABSTRACT

This study strives to create a real and practical solution to the optimization of digital marketing campaigns in a mobile games business. The current process treats information in a binary way and factors as independent from each other. Through a Fractional Factorial Experiment 2^{k-p} , the current model was compared with the proposed one with the intent of defining which one is more advantageous to the company.

In this project, concepts of Design of Experiments are used to elaborate a testing model for marketing videos with iterations of elements, that can be put in practice by the company. Firstly, the relevant characteristics of a Fractional Factorial Experiment 2^{k-p} were detailed. The mobile games market, the company positioning in it and the current process and its particularities were, then, exposed. Later, two equivalent projects were structured and modeled for analyzing one video with the same alterations made. The first one is exactly how the company does this kind of analysis nowadays, and the second one follows the proposed model. Both projects were executed and their impacts in the final results were analyzed and compared.

As for the results, the proposed model shows a superior performance, less dependent of intuition and unproven previous knowledges. Despite its higher production and testing costs, the proposed model had sufficiently better results to compensate this difference. By using the Design of Experiments based process, the company can optimize its campaigns performances, reducing considerably the operation costs and raising the general company's results.

Key Words: Design of Experiments; Fractional Factorial Experiment; Digital Marketing; Mobile Games

LISTA DE FIGURAS

Figura 1 - Estático do jogo Zooba, da Wildlife Studios.....	33
Figura 2 - Funil de conversão.....	35
Figura 3 - Indicadores do processo de conversão	37
Figura 4 - Iterações de personagem testadas em estático para o jogo Sniper 3D, da Wildlife Studios.....	41
Figura 5 – Exemplo do estilo cartoon de animação do jogo Castle Crush, da Wildlife Studios	43
Figura 6 - Exemplo de cena com apelo gráfico produzida pela célula de Motion 3D. Promoção de upgrade do jogo Tennis Clash, da Wildlife Studios	43
Figura 7 - Exemplo de animação com apelo realista e grande complexidade do jogo Battle Tanks, da Wildlife Studios.....	44
Figura 8 - Exemplo de animação bidimensional produzida em Motion 2D do jogo Math Learner, da Wildlife Studios	45
Figura 9 - Exemplo de alteração de fator binário em video do jogo Castle Crush, da Wildlife Studios.....	46
Figura 10 - Captura de tela do jogo Castle Crush, da Wildlife Studios	51
Figura 11 - Cartas dos personagens presentes no vídeo estudado: Warrior, Skeleton Swarm e Mud Elemental	52
Figura 12 - Storyboard do vídeo estudado com principais acontecimentos	53
Figura 13 - Cartela aplicada ao vídeo base	54
Figura 14 - Iteração de alteração do posicionamento dos esqueletos.....	56
Figura 15 - Iteração de alteração do cenário	56
Figura 16 - Iteração de adição de HUD.....	58
Figura 17 - Iteração de adição de elementos gráficos de sinalização do jogador	59
Figura 18 - Iteração de adição de cena com destaque para a carta.....	60
Figura 19 - Iteração de inclusão de cena de upgrade do cavaleiro	61
Figura 20 - Iteração de inclusão de cena de upgrade do monstro de lama	62
Figura 21 - Iteração de alteração da cartela	63
Figura 22 - IPM ratio por versão - Modelo Atual	74

Figura 23 – IPM ratio por versão - Experimento	78
Figura 24 - Variação de IPM em relação ao controle (vídeo 0) - Modelo Atual	79
Figura 25 - Resultados finais das versões consequentes da análise do Modelo Atual	91
Figura 26 - Resultados finais das versões consequentes da análise do primeiro conjunto de interações do Experimento	92
Figura 27 - Resultados finais das versões consequentes da análise do segundo conjunto de interações do Experimento	94

LISTA DE QUADROS

Quadro 1 - Detalhamento de experimentos em um delineamento fatorial 2^3	24
Quadro 2 - Blocagem de um experimento fatorial 2^3	28
Quadro 3 - Exemplo de experimento fatorial fracionado 2^{4-1}	29
Quadro 4 - Sinais para matriz ortogonal de 16 rodadas (Fonte: BOX, HUNTER e HUNTER, 2005)	64
Quadro 5 - Correspondência entre os fatores do modelo utilizado e as iterações propostas para o vídeo base	65
Quadro 6 - Sinais para as 16 rodadas do experimento delineado	66
Quadro 7 - Correspondência entre fatores fictícios do modelo e iterações propostas confundidas	67
Quadro 8 - Sinais para as 16 rodadas do experimento delineado com definição de blocagem feita a partir da interação dos fatores fictícios abcd	69
Quadro 9 - Versões produzidas para o Modelo Atual	71
Quadro 10 - Resultados do Teste 1 - Modelo Atual	73
Quadro 11 - Resultados do Teste 2 - Modelo Atual	73
Quadro 12 - IPM total e IPM ratio para os vídeos do Modelo Atual	74
Quadro 13 - Resultados do Teste 3 - Experimento, Bloco I	75
Quadro 14 - Resultados do Teste 4 - Experimento, Bloco I	75
Quadro 15 - Resultados do Teste 5 - Experimento, Bloco II	76
Quadro 16 - Resultados do Teste 6 - Experimento, Bloco II	76
Quadro 17 - IPM total e IPM ratio para os vídeos do Experimento	77
Quadro 18 - Efeitos individuais calculados para o Experimento	81
Quadro 19 - Efeitos de interações de dois fatores calculados para o Experimento .	82
Quadro 20 - IPM por rodada e variâncias	83
Quadro 21 - Variâncias dos IPM ratios para cada versão	84
Quadro 22 - Efeitos e erros finais calculados para o Experimento	85
Quadro 23 - Versões produzidas para os testes finais	89
Quadro 24 - IPM e IPM ratio obtidos nos testes finais	89

LISTA DE ABREVIATURAS E SIGLAS

PC – *Personal Computer*

App – *Application*

CTA – *Call-To-Action*

CPM – *Cost Per Mille*

CTR – *Click Through Rate*

IR – *Install Rate*

LTV – *Lifetime Value*

ROI – *Return On Investment*

CPI – *Cost Per Install*

IPM – *Install Per Mille*

CPA – *Cost Per Action*

Adnet – *advertising network*

IAP – *In-App Purchases*

SUMÁRIO

1. INTRODUÇÃO.....	21
1.1. CONTEXTO.....	21
1.2. OBJETIVOS E ESTRUTURA.....	22
2. ESTATÍSTICA EXPERIMENTAL	23
2.1. EXPERIMENTO FATORIAL 2^k	24
2.1.1. Cálculo dos Efeitos do Experimento Fatorial 2^k	25
2.1.2. Cálculo dos Erros do Experimento Fatorial 2^k	26
2.1.3. Blocagem de experimentos	27
2.1.4. Experimento Fatorial Fracionado 2^{k-p}	28
2.2. VANTAGENS	29
3. O MERCADO DE <i>MOBILE GAMES</i>.....	31
3.1. O MARKETING DIGITAL PARA A INDÚSTRIA DE <i>MOBILE GAMES</i>	32
3.1.1. Criativos e Anúncios	32
3.1.3. Redes Sociais e <i>Adnets</i>	38
3.2. A EMPRESA.....	40
3.2.1. Processo Atual.....	41
4. PLANEJAMENTO	51
4.1. O VÍDEO BASE	52
4.2. ITERAÇÕES.....	55
4.2.1. Alteração do posicionamento dos esqueletos	55
4.2.2. Alteração do cenário	56
4.2.3. Adição de <i>HUD</i>	57
4.2.4. Adição de elementos gráficos de sinalização do jogador	58

4.2.5.	Inclusão de cena com destaque para a carta	59
4.2.6.	Inclusão de cena de <i>upgrade</i> do cavaleiro.....	60
4.2.7.	Inclusão de cena de <i>upgrade</i> do monstro de lama	62
4.2.8.	Alteração da cartela	63
4.3.	MODELAGEM.....	64
4.3.1.	Blocagem	68
5.	EXECUÇÃO E RESULTADOS	71
5.1.	TESTAGEM	72
5.1.1.	Testes e resultados para o Modelo Atual.....	72
5.1.2.	Testes e resultados para o Experimento.....	75
6.	ANÁLISE.....	79
6.1.	ANÁLISE DOS RESULTADOS DO MODELO ATUAL	79
6.2.	ANÁLISE DOS RESULTADOS DO EXPERIMENTO	81
6.2.1.	Efeitos de Fatores Individuais	81
6.2.2.	Efeitos de Interações de Dois Fatores	82
6.2.3.	Variância e Desvio Padrão.....	82
6.2.4.	Análise dos resultados obtidos para o Experimento	85
6.2.5.	Resultados dos testes finais.....	88
6.2.6.	Comparação final dos modelos.....	89
7.	CONCLUSÕES E PRÓXIMOS PASSOS	97
8.	REFERÊNCIAS.....	99

1. INTRODUÇÃO

1.1. CONTEXTO

O mercado de jogos digitais é atualmente o maior setor da indústria do entretenimento. Em 2018, os *games* se tornaram mais lucrativos que a indústria do cinema e da música juntas, movimentando mundialmente US\$ 137 bilhões (GAGLIONI, 2019). Dentro desse mercado, o de jogos digitais para celular (*mobile games*¹) tornou-se nos últimos anos o mais expressivo, correspondendo hoje a mais de 50% de toda a indústria dos *video games* e com o expressivo crescimento de 26% ao ano (PRNEWswire, 2019). Dentro desse contexto, a Wildlife Studios é uma produtora de *mobile games* brasileira que se firmou como uma das maiores do mundo no final de 2019 ao ser avaliada em 1,3 bilhão de dólares (GRANATO, 2020).

O modelo de negócio da Wildlife tem o marketing digital como elemento central para distribuição de jogos. O marketing de *mobile games* baseia-se, principalmente, nos chamados criativos, que são vídeos curtos, imagens e anúncios jogáveis (*playables*), que são produzidos internamente e divulgados em diferentes aplicativos. Esse setor é de importância estratégica e representa atualmente o principal gargalo para a disseminação e consequente retorno financeiro dos jogos produzidos. Anualmente, os criativos, que promovem os jogos da empresa, são reproduzidos bilhões de vezes, em todos os países do mundo. Por isso a área de *Creatives Intelligence* (*Creatives Intel* ou Inteligência dos Criativos) tem como principal função otimizar o desempenho desses criativos e o retorno das campanhas publicitárias.

A área de *Creatives Intel* é muito recente, tanto na empresa como no mundo, e possui boa parte de seus métodos e processos ainda em desenvolvimento. Dessa forma, sabe-se que sua atuação ainda não é a ideal e procura-se constantemente aprimorar seu desempenho para atingir os objetivos da empresa.

¹ a partir deste ponto, os jogos digitais móveis serão denominados de *mobile games*, seguindo a prática de mercado

1.2. OBJETIVOS E ESTRUTURA

O objetivo deste trabalho é o aprimoramento dos testes de desempenho de criativos, uma das principais etapas do seu processo de produção. O teste de diferentes versões de vídeos, imagens e *playables* é o foco da área de *Creatives Intel*, e é a base da otimização do desempenho desses anúncios, ou seja, a obtenção de mais usuários (mais *downloads*) por um custo de publicidade menor. Esse processo, apesar de extremamente importante, ainda é realizado de forma pouco otimizada, principalmente por causa de sua grande complexidade e da alta demanda de produção de criativos, que reduz o tempo dedicado ao estudo da eficiência das campanhas publicitárias.

Com isso em mente, esse trabalho pretende apresentar uma solução alternativa para essa etapa do processo, e a realização de um experimento para comparação de resultados com o atual modelo. Para isso, primeiro fez-se uma revisão teórica da estatística experimental e do modelo utilizado e uma introdução a respeito do mercado de *mobile* games, resumindo-se as principais informações sobre ele, o *marketing digital* nessa indústria, a empresa em que se baseia o estudo (Wildlife Studios). Depois, o experimento para comparação com o processo atual será desenhado e planejado para, finalmente, ser executado. Após sua execução, os resultados serão analisados e estudados para que conclusões sejam feitas.

2. ESTATÍSTICA EXPERIMENTAL

A estatística experimental é o ramo da matemática que estuda dados obtidos através de experimentos (ou ensaios), que são pesquisas planejadas para confirmação ou refutação de hipóteses ou outros resultados obtidos anteriormente (FERREIRA, 2011).

Em um experimento delineado, projeta-se ensaios em que se faz variações propositalmente nas variáveis de entrada (fatores) de um processo de forma que sua influência nas variáveis de saída seja calculada. Assim, pode-se determinar quais fatores afetam a variável de saída e qual a magnitude de seu impacto no resultado. O planejamento formal de um experimento visando se obter conhecimentos a respeito de determinados fatores é chamado de delineamento experimental (CALEGARE, 2009).

Assim, o Delineamento de Experimentos é o processo de planejamento, desenho e análise de um experimento construído para que conclusões claras e objetivas sejam obtidas de maneira eficiente e efetiva. Segundo Antony (2014), o Delineamento de Experimentos abrange os Delineamentos de Peneiramento (*Screening Design*), também conhecido como Experimento de Plackett-Burman, o Experimento Fatorial Completo e o Experimento Fatorial Fracionado.

Os Delineamentos de Peneiramento são planejamentos simples em que se desconsidera a interação entre os fatores de entrada nos resultados do processo. Têm como objetivo selecionar (peneirar) os múltiplos fatores que podem afetar um processo para que os que possuem maior impacto possam ser estudados mais a fundo.

Em um Experimento Fatorial Completo, realiza-se todas as combinações possíveis de variações dos valores dos fatores que afetam os resultados de um processo. Nesse delineamento, pode-se obter conclusões a respeito das interações entre fatores, observando-se as combinações em que ambos variam.

Já o Experimento Fatorial Fracionado permite que conclusões sejam obtidas a partir da realização de apenas uma fração das combinações entre os possíveis valores de cada fator. Nesses casos, alguns efeitos de fatores ou de interações entre fatores são confundidos, limitando os aprendizados diretos, mas possibilitando que menos testes sejam realizados.

O experimento a ser realizado nesse projeto se baseará em um experimento fatorial fracionado 2^{k-p} estruturado a partir da definição de Box, Hunter e Hunter (2005).

2.1. EXPERIMENTO FATORIAL 2^k

De acordo com Box, Hunter e Hunter (2005), um experimento fatorial é aquele em que o número de versões para cada fator (variável) é fixo e realiza-se todas as combinações possíveis entre eles. Segundo Montgomery e Runger (2003), um experimento fatorial é a única maneira de se descobrir interações de impacto entre variáveis, sendo sempre preferível à variação de um fator por vez na análise do impacto de múltiplos fatores em um resultado.

O experimento fatorial em que existem apenas duas variações para cada fator é chamado de experimento fatorial 2^k . No Quadro 1, temos o exemplo de um detalhamento de um experimento fatorial 2^3 , em que existem três fatores que variam entre dois valores, levando a 8 combinações possíveis entre fatores, que representam os experimentos (numerados de 1 a 8). No exemplo, para os fatores hipotéticos A, B e C, os sinais de mais (+) e menos (-) representam valores que cada um deles pode assumir. Existem ainda as interações entre os fatores em cada um dos experimentos, representados nas quatro últimas colunas do Quadro 1. Dessa forma, AB representa a interação entre os fatores A e B e assim por diante para as colunas restantes.

Quadro 1 - Detalhamento de experimentos em um delineamento fatorial 2^3

Experimentos	A	B	C	AB	AC	BC	ABC
1	-	-	-	+	+	+	-
2	+	-	-	-	-	+	+
3	-	+	-	-	+	-	+
4	+	+	-	+	-	-	-
5	-	-	+	+	-	-	+
6	+	-	+	-	+	-	-
7	-	+	+	-	-	+	-
8	+	+	+	+	+	+	+

Nota-se que o sinal utilizado nas quatro últimas colunas do Quadro 1 é proveniente da multiplicação entre os sinais de cada fator naquele experimento. Dessa forma, um quadro de sinais equivalente é facilmente obtido para qualquer número de fatores (k).

Para cada coluna do Quadro 1, percebe-se que o número de sinais positivos e negativos é sempre igual a quatro. Além disso, um sinal (positivo ou negativo) é encontrado simultaneamente em duas colunas quaisquer na mesma linha (experimento) exatamente duas vezes. Por exemplo, para as colunas B e AC, dois sinais positivos são encontrados apenas nos experimentos 3 e 8 e dois sinais negativos apenas nos experimentos 2 e 5. Essa característica de equilíbrio recebe o nome de ortogonalidade e possui papel fundamental para os experimentos fatoriais. A ortogonalidade garante que as estimativas para cada efeito não sejam afetadas pelos sinais e valores dos demais efeitos.

2.1.1. Cálculo dos Efeitos do Experimento Fatorial 2^k

A partir do quadro de sinais (Quadro 1), tanto Box, Hunter e Hunter (2005) quanto Montgomery e Runger (2003) sugerem um método simples de cálculo dos efeitos de cada fator. O efeito do fator A, por exemplo, é o impacto que a alteração deste fator entre seus dois valores possíveis causa no resultado final. Para se calcular isso, deve-se subtrair a média dos resultados obtidos com o sinal negativo para o fator da média dos resultados obtidos com o sinal positivo. Para o fator A, portanto, o efeito calcula-se da seguinte forma:

$$l_A = \frac{y_2 + y_4 + y_6 + y_8}{8} - \frac{y_1 + y_3 + y_5 + y_7}{8}$$

Onde l_A é o efeito do fator A e y_n , o resultado obtido pelo experimento n. Na primeira fração inclui-se todas as versões em que o sinal para A foi positivo (todas as pares) e na segunda, as versões em que o sinal para A foi negativo (todas as ímpares). O cálculo é realizado de maneira análoga para todos os efeitos, inclusive os resultantes das interações entre fatores.

2.1.2. Cálculo dos Erros do Experimento Fatorial 2^k

Além dos efeitos de cada fator e interação entre fatores, é interessante para um Experimento Fatorial 2^k que se saiba o erro padrão desses valores. Ele é utilizado como parâmetro para se verificar quais efeitos possuem impacto significativo no resultado avaliado. Os que possuem suas variações dentro do intervalo de mais ou menos o desvio padrão, não são considerados relevantes. Box, Hunter e Hunter (2005) explicam que o erro padrão deve ser a raiz quadrada da variância de um efeito que, por sua vez, é uma fração da variância média dos experimentos.

A variância de um efeito deve ser uma fração da variância média dos experimentos de acordo com a forma que o efeito é calculado. No caso exemplificado pelo Quadro 1, os efeitos são obtidos pela diferença da média entre dois grupos de quatro fatores e, portanto, a variância média dos experimentos deve ser multiplicada por $\left(\frac{1}{4} + \frac{1}{4}\right)$. De forma genérica, em um Experimento Fatorial 2^k a variância de um efeito pode, portanto, ser calculada da seguinte forma:

$$\sigma_{efeito}^2 = \left(\frac{1}{k/2} + \frac{1}{k/2}\right) \bar{\sigma}_{experimentos}^2$$

Onde σ_{efeito}^2 é a variância de cada efeito, k é o número de fatores sendo alterados e $\bar{\sigma}_{experimentos}^2$ é a variância média dos experimentos, que pode ser calculada como uma média simples de cada uma das variâncias obtidas.

A variância de um experimento deve ser obtida ou aproximada de acordo com a forma como o experimento é conduzido. Nos casos em que os experimentos são realizados com repetições, ela é a variância entre os valores obtidos em cada repetição para cada experimento. Especificamente para experimentos com apenas duas repetições, o cálculo da variância pode ser simplificado para:

$$\sigma_{experimento}^2 = \frac{(y_a - y_b)^2}{2}$$

Onde y_a e y_b são os resultados obtidos em cada uma das repetições realizadas.

A variância de uma razão A/B , pode ser aproximada, segundo Seltman (2018), como:

$$\sigma_{A/B}^2 = \frac{A^2}{B^2} \left[\frac{\sigma_A^2}{A^2} + \frac{\sigma_B^2}{B^2} - 2 \frac{cov(A, B)}{AB} \right]$$

2.1.3. Blocagem de experimentos

Em um experimento fatorial, nem sempre se consegue fazer todos os experimentos com os mesmos recursos. Para se neutralizar os efeitos da diferença entre blocos (ou lotes), usa-se o conceito de blocagem. Procura-se agrupar os experimentos em lotes visando reduzir a importância dos fatores ou interações afetados por ele. No exemplo de um experimento fatorial 2^3 , como o observado no Quadro 1, em que se pode apenas realizar quatro rodadas de experimento em cada lote, os experimentos são agrupados pela interação de maior ordem possível, ou seja, a interação de três fatores (ABC). Isso é feito porque, em geral, quanto maior a ordem da interação, menor o seu efeito nos resultados. Dessa forma, realiza-se deliberadamente o confundimento do efeito causado pela interação ABC com o efeito da blocagem. Assim, não se é possível chegar a conclusões concretas a respeito dessa interação específica, mas todas as outras análises mantêm-se intactas devido à ortogonalidade dos experimentos presentes. O Quadro 2 mostra como a blocagem é realizada, agrupando-se os experimentos em que a interação ABC assume o mesmo valor (+ ou -) entre os lotes I e II.

Em uma blocagem de experimento fatorial, nem sempre se consegue confundir os lotes com as interações de maior ordem. Nesse caso, busca-se o confundimento com interações das quais se espera a menor interferência nos resultados, que são, normalmente, as de maior ordem possível.

Quadro 2 - Blocagem de um experimento fatorial 2^3

Experimentos	A	B	C	AB	AC	BC	ABC	LOTE
1	-	-	-	+	+	+	-	I
2	+	-	-	-	-	+	+	II
3	-	+	-	-	+	-	+	II
4	+	+	-	+	-	-	-	I
5	-	-	+	+	-	-	+	II
6	+	-	+	-	+	-	-	I
7	-	+	+	-	-	+	-	I
8	+	+	+	+	+	+	+	II

2.1.4. Experimento Fatorial Fracionado 2^{k-p}

Os experimentos fatoriais, ao exigirem que todas as combinações entre fatores sejam testadas, demandam um elevado número de amostras e de execuções. Em muitos processos isso torna o experimento muito custoso ou demorado. Para casos em que os recursos ou tempo não são suficientes, os experimentos fatoriais fracionados, de acordo com Box, Hunter e Hunter (2005) e Montgomery e Runger (2003), oferecem uma alternativa. Neles, apenas uma fração do total de combinações dos fatores é produzida e testada. Um experimento fatorial fracionado 2^{k-p} é aquele em que k fatores permitem uma variação entre dois valores, e apenas uma fração de $(1/2)^p$ dos experimentos é realizada.

Para tornar os experimentos fatoriais fracionados possíveis e seus resultados aproveitáveis, utiliza-se o mecanismo de confundimento (semelhante ao da blocagem) em que algumas interações e fatores são deliberadamente coincidentados em todos os experimentos. Isso limita conclusões sobre alguns níveis de interações, mas possibilita que diversas análises sejam realizadas sem a necessidade de um número de execuções muito elevado. Em um experimento fatorial 2^{4-1} , por exemplo, utiliza-se como base o quadro de sinais de um experimento fatorial 2^3 (como o Quadro 1) com fatores fictícios. O Quadro 3 mostra um exemplo em que os fatores fictícios a , b e c , assim como suas interações, são representados. A partir do quadro de sinais de

fatores fictícios, utilizado como base, confunde-se os fatores reais, da mesma forma que na blocagem, com os fatores fictícios e suas interações. Usualmente, confunde-se, primeiro, com todos os fatores fictícios individuais e, depois, com as interações do maior número de fatores possível. No Quadro 3, além dos sinais para os fatores fictícios a, b e c, existe um exemplo de um experimento fatorial fracionado 2^{4-1} , representado pelos fatores em letra maiúscula A, B, C e D. Os fatores A, B e C são confundidos com os fatores fictícios a, b e c de forma intuitiva. Já o fator D é confundido com a interação abc. Para isso, utiliza-se para o fator D o sinal (+ ou -) correspondente à multiplicação dos sinais de A (a), B (b) e C (c).

Quadro 3 - Exemplo de experimento fatorial fracionado 2^{4-1}

Experimentos	a	b	c	ab	ac	bc	abc
	A	B	C				D
1	-	-	-	+	+	+	-
2	+	-	-	-	-	+	+
3	-	+	-	-	+	-	+
4	+	+	-	+	-	-	-
5	-	-	+	+	-	-	+
6	+	-	+	-	+	-	-
7	-	+	+	-	-	+	-
8	+	+	+	+	+	+	+

O confundimento para um experimento fatorial fracionado pode ser simples para casos de poucos fatores, mas torna-se muito complexo conforme o número aumenta. A literatura oferece tabelas em que os confundimentos para diversas variações de k e p em experimentos fatoriais fracionados 2^{k-p} estão já estipulados.

2.2. VANTAGENS

O delineamento de um experimento fatorial permite que se identifique a interação entre diferentes fatores, isso é, a correlação entre a variação de fatores. Isso não é possível quando se testa um fator de cada vez. Além disso, esse modelo permite

que se atinja níveis de confiança nos resultados maiores com a realização de menos experimentos o que, em muitos casos, significa uma grande economia de recursos.

3. O MERCADO DE *MOBILE GAMES*

A indústria global de jogos digitais, incluindo *mobile*, PC, consoles, vídeos relacionados a jogos e realidade aumentada teve uma receita total de US\$120,1 bi em 2019 (crescimento de 4% em relação a 2018). Desse total, US\$64,4 bi (53,62%) foram de *mobile games* (TAKAHASHI, 2020). É inegável que esse segmento de mercado representa uma parte importante não só da indústria do entretenimento como também da economia mundial. Dentre os principais motivos para esse protagonismo, está o aumento constante do uso de *smartphones*. O número de usuários de *smartphones* cresce anualmente e, em 2017, já passava os 2,3 bilhões (RUTZ, ARAVINDAKSHAN e RUBEL, 2019).

Além do enorme e crescente número de potenciais clientes, o mercado ainda possui a vantagem de que seus produtos são 100% virtuais. Dessa forma, um jogo desenvolvido para celular pode chegar ao mesmo tempo em todos os países do mundo, sem os custos logísticos e desembaraço alfandegário dos produtos físicos em mercados tradicionais.

O tamanho do mercado e a facilidade de distribuição são os grandes atrativos desse segmento, mas representam também alguns dos principais obstáculos enfrentados. A concorrência é expressiva: o número de aplicativos na Google Play Store, uma das maiores lojas de aplicativos no mundo, era de 2,87 milhões em março de 2020, de acordo com a própria loja.

Além disso, o crescimento orgânico do número de *downloads* de um *app*, ou seja, o número de pessoas que baixam espontaneamente um aplicativo, sem serem influenciadas por um anúncio, isto é, que encontram o aplicativo por indicação ou por busca, é muito baixo. Embora esse número varie de acordo com o aplicativo ou a época, na maior parte dos casos é necessário um grande investimento em marketing digital para o sucesso de um *mobile game*. Tal dispêndio não é possível para pequenos estúdios e desenvolvedores autônomos, tornando a competição com grandes produtoras difícil e desigual. O mercado de *mobile games* é, portanto extremamente promissor, mas extremamente competitivo, e o marketing digital possui um papel fundamental para o êxito das empresas no ramo.

3.1. O MARKETING DIGITAL PARA A INDÚSTRIA DE *MOBILE GAMES*

O marketing digital é, para a indústria de *mobile games*, a principal fonte de obtenção de clientes. Por trabalhar com produtos virtuais e devido ao enorme número de opções de oferta, encontrou-se nesse modelo de anúncio a melhor maneira de se atingir pessoas em todo o mundo.

3.1.1. Criativos e Anúncios

Criativo é o nome que se dá no ramo do marketing a um produto artístico utilizado para divulgação de um produto. Os criativos feitos para jogos de celular concentram-se em três tipos: estáticos, vídeos e *playables*.

- Estáticos são imagens estáticas, que geralmente possuem um botão com uma mensagem que convida a pessoa a clicar, conhecida como *call-to-action*, ou CTA. Exemplos de CTA comumente usados são “*Play now!*”, ou “Jogue agora!” e “Download for free!”, ou “Baixe gratuitamente!”. A Figura 1 é um exemplo de estático do jogo *Zooba – Zoo Battle Arena*, lançado em 2019 pela Wildlife Studios.
- Vídeos são a forma mais utilizada pelos *games* para divulgação. Eles podem mostrar capturas de trechos reais do jogo (*gameplays*), animações que representem os personagens ou situações do jogo, ou uma combinação dos dois.
- *Playables* são a tecnologia mais recente no universo de marketing de *mobile games*. Eles são anúncios jogáveis, ou seja, trechos ou adaptações do jogo extremamente curtos e simplificados para possibilitar a experiência do jogador antes do *download*.



Figura 1 - Estático do jogo Zooba, da Wildlife Studios

Um anúncio é aquilo que é mostrado ao consumidor final e pode ser um criativo ou uma combinação de criativos. Os anúncios mais comuns são: estáticos, vídeos e *playables* independentes, vídeo seguido de um estático, *playable* seguido de um estático e vídeo seguido de um *playable*.

Existem dois tipos de anúncios digitais: *interstitial ads* e *rewarded ads*. *Interstitial ads* são anúncios mostrados sem que haja uma recompensa para o consumidor. Eles aparecem na *timeline* ou *stories* de redes sociais e são mostrados entre fases de jogos. A principal característica desse tipo de anúncio é que ele não precisa ser visto até o fim, após cinco segundos de exibição, é permitido ao usuário fechá-lo. Já um *rewarded ad* é um anúncio que o jogador opta por assistir para conseguir uma recompensa, eles são muito comuns em jogos para se obter uma vida extra ou algum item consumível. Esse tipo de anúncio deve ser assistido até o final para que se consiga obter a recompensa.

3.1.2. Campanhas

Uma campanha é um conjunto de anúncios que é mostrado para um determinado público por um determinado período de tempo. Um público pode ser estratificado por faixa etária, gênero, localização geográfica, atividade em redes sociais e histórico de downloads, entre outros. Dessa forma, a definição das campanhas permite a personalização e direcionamento preciso dos anúncios. A duração de uma campanha é o intervalo de tempo em que ela fica no ar e pode ser definida previamente pelo anunciante ou não. Para aumentar seu desempenho, uma campanha possui diversos anúncios, mostrados a um público específico em um mesmo período de tempo, e, portanto, os anúncios competem entre si.

3.1.2.1. *Funil de conversão*

Além do público e da duração, outros fatores caracterizam uma campanha. Eles podem ser definidos pelo anunciante, obtidos como resultados ou otimizados automaticamente, dependendo do tipo de campanha. Os principais fatores são: orçamento, impressões, *clicks*, instalações e receita.

Orçamento, ou *budget*, é o custo de uma campanha. Ele pode ser definido pelo anunciante previamente ou ser consequência da performance obtida por ela. O orçamento pode ser definido por anúncio, por campanha ou por dia, de maneira a limitar o número de vezes que cada anúncio é mostrado em cada dia.

Impressão é o nome que se dá a cada vez que um anúncio é exibido a uma pessoa e, portanto, o número total de impressões de uma campanha é importante para se analisar seus resultados. Usualmente, para que os resultados de uma campanha sejam confiáveis, define-se um número mínimo de impressões necessárias por anúncio. O número de impressões em campanhas comuns é obtido como resultado e varia entre anúncios de acordo com os outros parâmetros definidos e das características de cada criativo usado. Porém, em determinados tipos de campanha, como as campanhas teste, define-se que todos os anúncios devem ter quantidades iguais de impressões.

O número de *clicks* é a quantidade de pessoas que clicam no anúncio que recebeu, sendo direcionadas para a página da loja de aplicativos em que podem fazer ou não o download. O número de *clicks* é uma fração do número de impressões e, em

campanhas comuns, é obtido como resultado e varia entre anúncios de acordo com os outros parâmetros definidos e das características de cada criativo usado.

Instalações são o objetivo final de uma campanha, e são o número de vezes em que uma pessoa que assistiu a um anúncio faz *download* do jogo anunciado em seguida. O número de instalações é uma fração do número de *clicks* e, em campanhas comuns, é obtido como resultado e varia entre anúncios de acordo com os outros parâmetros definidos e das características de cada criativo usado.

A receita de uma campanha é definida como o valor gasto por todos os usuários que baixam o jogo por causa dela, durante todo o tempo em que eles ou elas jogam. Por isso, a receita não é definida imediatamente e só pode ser conhecida após longos períodos de tempo, dependendo da média de tempo em que um jogador permanece jogando. Mesmo assim, ela é um fator importante para a definição de diversas outras métricas. O número de jogadores que fazem uma compra dentro do aplicativo é uma fração do número de jogadores que instalam o jogo.

Esses fatores formam o chamado funil de conversão (Figura 2), que representa a trajetória dos consumidores, desde a exibição do anúncio até eventuais compras feitas dentro do jogo. A forma de funil é definida por causa do fato de que cada etapa só é alcançada por uma fração das pessoas que chegaram na etapa anterior.

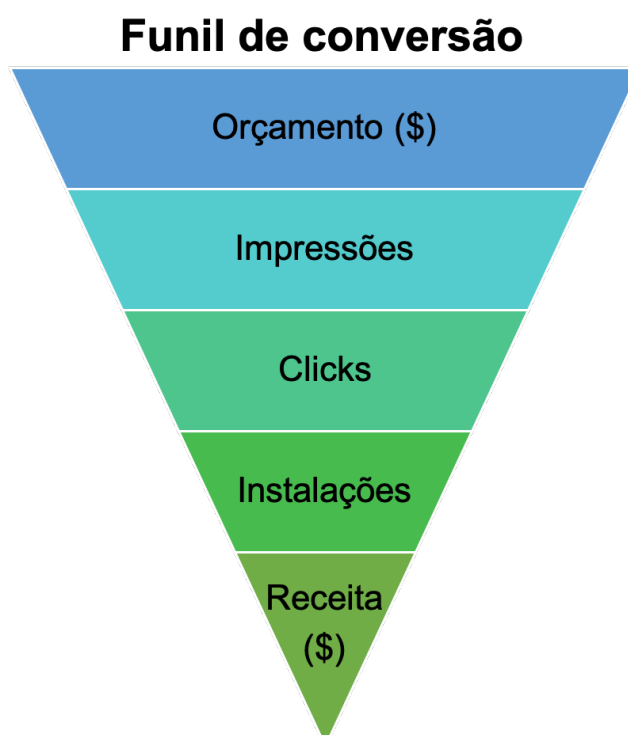


Figura 2 - Funil de conversão

3.1.2.2. Indicadores

A partir do funil de conversão, diversos indicadores são obtidos e utilizados como parâmetros e métricas das campanhas. Abaixo, cada um deles é explicado.

Cost Per Mille (CPM) é o valor pago a cada mil impressões. Ele é calculado dividindo-se o orçamento pela milésima parte do número de impressões. Utiliza-se a proporção do milhar devido ao baixo valor pago por impressão, permitindo-se que os números obtidos sejam mais amigáveis e significativos.

Click Through Rate (CTR) é a taxa de *clicks* em comparação ao volume de impressões, calcula-se dividindo o valor do primeiro pelo do segundo.

Install Rate (IR) é a taxa de instalação em comparação ao número de *clicks* e é a fração do primeiro em relação ao segundo.

Lifetime Value (LTV) é o valor médio arrecadado no *app* por jogador e calcula-se dividindo a receita do jogo pelo número de instalações.

Return On Investment (ROI) é a taxa de retorno obtida em relação ao valor gasto e é obtida dividindo-se a receita pelo orçamento. É importante ressaltar que este é o ROI de marketing e considera apenas os custos das campanhas e a receita gerada pelos usuários obtidos por meio delas.

Cost Per Install (CPI) é o custo por instalação, ou seja, o valor gasto dividido pelo número de instalações obtidas.

Install Per Mille (IPM) é o número de instalações obtido a cada mil impressões, calcula-se dividindo o número de instalações pela milésima parte do número de impressões. Utiliza-se a proporção do milhar devido ao elevado número de impressões necessária para se obter um *download*, tornando os indicadores mais amigáveis.

A Figura 3 ilustra como os principais indicadores do processo de conversão relacionam-se com o Funil.

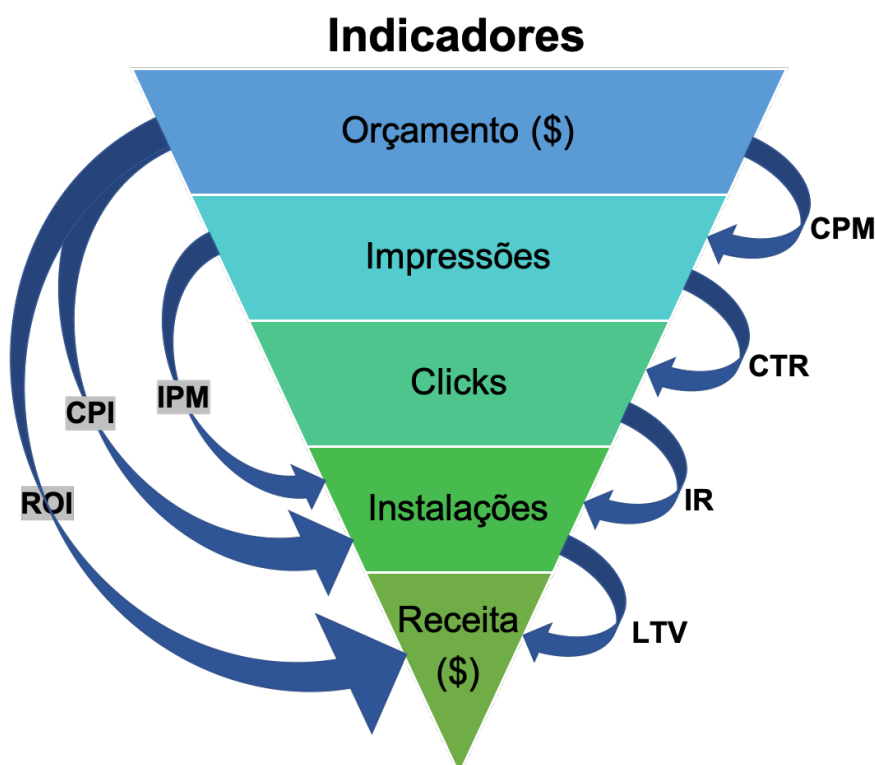


Figura 3 - Indicadores do processo de conversão

Além desses, existe um outro importante indicador não relacionado diretamente ao modelo usual do funil de conversão, mas que também é muito utilizado para definir e avaliar campanhas. Chama-se CPA ou *Cost Per Action* o custo médio de aquisição de usuários por ação realizada dentro do jogo. Essa ação realizada é definida pelo anunciante e pode variar de acordo com os interesses estratégicos de cada *app* e normalmente correspondem a uma compra feita ou um valor mínimo gasto. Esse indicador é obtido dividindo-se o orçamento pelo número de vezes que a ação correspondente é realizada pelos usuários obtidos pela campanha.

As campanhas são controladas por algoritmos que as otimizam de acordo com os resultados esperados. A otimização é feita equilibrando-se a quantidade de vezes que cada anúncio dentro da campanha é enviado para impressão, de acordo com os resultados observados. Por isso, o algoritmo precisa de um tempo de aprendizado para que funcione corretamente. Existem diversos tipos de otimização de campanha, sendo os mais comuns a otimização por CPI, por IPM e por CPA. As campanhas otimizadas por CPA possuem um tempo de aprendizado mais alto, pois dependem de uma ação dentro do jogo realizada por um usuário obtido pela campanha. Por isso, ela só é utilizada em campanhas mais longas ou permanentes. As outras duas (CPI e

IPM) apresentam resultados mais imediatos, porém menos precisos para se conhecer a vantagem que cada anúncio apresenta em relação aos outros.

3.1.3. Redes Sociais e *Adnets*

As campanhas são mostradas ao público em aplicativos diversos. Os dois principais tipos de aplicativos em que as pessoas são impactadas por anúncios são redes sociais e os próprios jogos. As redes sociais possuem diferentes locais em que um anúncio pode ser impresso. No Facebook, por exemplo, um usuário pode ver um anúncio no *feed* de notícias, nos *stories* ou na forma de *banners* (posicionados nas laterais ou na região inferior da tela). Assim como o Facebook, outras redes sociais possuem espaços disponíveis para anúncios, como Instagram, Youtube, Twitter, Pinterest, Snapchat e Tiktok. Nas redes sociais, os anúncios são sempre do tipo *interstitial*, enquanto nos aplicativos de jogos, os anúncios aparecem tanto na forma *interstitial* como *rewarded*.

A principal diferença entre anunciar em uma rede social ou em um outro aplicativo de jogo é a forma como a negociação é realizada. Para se anunciar em redes sociais, compra-se diretamente da empresa responsável pelo espaço para o anúncio, assim, para anúncios no Facebook ou no Instagram, negocia-se com o Facebook, empresa que possui as duas redes, ou para se anunciar no Youtube, deve-se comprar espaço com o Google. Já no caso de anúncios em aplicativos de jogos, existem intermediários, as chamadas *advertising networks*, ou simplesmente *adnets*.

As *adnets* compram o espaço publicitário de aplicativos e o revendem para outros aplicativos. Dessa forma, é muito comum ver anúncios de um jogo dentro de um jogo concorrente. Atualmente todos os tipos de empresa fazem marketing digital e, por isso, nesses espaços observa-se não apenas anúncios de outros jogos, mas também dos mais diversos serviços digitais.

3.1.3.1. Leilão de anúncios

O processo de compra e venda de espaços para anúncios é muito semelhante entre redes sociais e *adnets*, sendo a única diferença a presença de um participante e uma etapa a mais no segundo caso. Esses espaços são comprados e vendidos por meio de um leilão constante e dinâmico entre anunciantes e redes (sociais ou *adnets*).

Quando se anuncia por meio de uma *adnet*, ela compra do aplicativo em que se vai anunciar o espaço para o anúncio, pagando por impressão feita, e revende esse espaço para os anunciantes, que pagam por instalação obtida. Dessa forma, tanto para a *adnet* quanto para o anunciante, os melhores anúncios são aqueles que possuem um IPM elevado. Para a primeira, isso garante um número mais baixo de impressões (custo menor) com um número elevado de instalações (receita maior). Para o anunciante, como se paga por instalação, o número de impressões feitas não é valioso, mas um IPM elevado garante que se consiga instalações mais rapidamente. Além disso, a forma como as compras e vendas são feitas tornam o IPM um indicador ainda mais relevante.

Para cada jogador ou usuário que vai receber um anúncio, é feito um pequeno leilão entre a *adnet* e todos os anunciantes, que acontece num período de tempo menor do que um segundo, e é realizado milhares de vezes a todo momento, por causa do alto número de pessoas constantemente acessando aplicativos e do baixo tempo de resposta esperado para que o usuário não precise esperar para ver o anúncio. Por isso, esses pequenos leilões são realizados por algoritmos montados por cada empresa anunciante, que são operados diariamente e em tempo real por equipes de *User Acquisition*. A competição entre anúncios é feita levando-se em conta o lance dos anunciantes, ou seja, o quanto eles estão dispostos a pagar por impressão e o IPM. O lance é feito pelos algoritmos e pelos operadores de campanhas, enquanto o IPM varia de acordo com o anúncio divulgado e o público para o qual ele é mostrado.

Um anúncio com IPM elevado permite ao anunciante fazer lances mais baixos, pois o volume de instalações será maior, com uma quantidade menor de impressões. Dessa forma, o custo da *adnet* por instalação é mais baixo e, portanto, ela aceita receber menos por elas. Já um anúncio com um IPM baixo exige do anunciante lances altos. Dessa forma, um IPM alto é favorável tanto para o anunciante quanto para a *adnet*. O leilão é feito equilibrando-se IPM e lance para os anúncios escolhidos e, assim, otimizando-se os resultados para a *adnet*.

3.2. A EMPRESA

A Wildlife Studios é uma das dez maiores produtoras de *mobile games* do mundo. Ela foi fundada em São Paulo em 2011 por dois irmãos formados em engenharia pela Escola Politécnica da Universidade de São Paulo. Atualmente possui seis escritórios em quatro países (Brasil, Estados Unidos, Argentina e Irlanda) e mais de 500 funcionários. A empresa possui em seu portfólio mais de 70 jogos e já acumulou dois bilhões de *downloads* (WILDLIFE STUDIOS, 2020). No final de 2019, após sua primeira rodada de investimentos, a empresa foi avaliada em 1,3 bilhão de dólares (GRANATO, 2020).

Hoje, a empresa trabalha apenas com o modelo de negócio *freemium*. Nesse modelo, os *apps* são baixados pelo usuário de forma gratuita, mas possuem funcionalidades limitadas, que possuem prazo para vencer ou que são oferecidas no modo *trial* para que o jogador decida se deseja pagar pelo produto. No modelo *freemium*, as principais fontes de renda para o desenvolvedor são compras dentro do *app* (IAP ou *in-app purchases*) e/ou a venda de espaço para anúncios. Em uma IAP, o jogador compra, dentro do *app*, itens virtuais que oferecem melhorias dentro do jogo, de jogabilidade ou apenas estética (TANG, 2019). Dessa forma, *mobile games* permitem monetizar o engajamento dos jogadores além da compra inicial, como normalmente acontece com outros tipos de *games* (RUTZ, ARAVINDAKSHAN e RUBEL, 2019). Segundo Takahashi (2020), 80% do dinheiro gasto em jogos digitais em 2019 foram em jogos gratuitos (*freemium*).

Um dos principais fatores limitantes de crescimento desse modelo é o número de jogadores, pois apenas uma fração deles acaba realizando alguma compra. Dessa forma, a aquisição de novos usuários é de extrema importância para a empresa e, por isso, o investimento em marketing é elevado. No setor dos *games* é muito comum que empresas terceirizem a produção de anúncios, mas na Wildlife existe uma equipe de artistas com essa tarefa.

Levando-se em consideração que o IPM é de extrema importância para a otimização de custos de campanhas e que a produção de conteúdos digitais é interna, o que permite grande flexibilidade e comunicação eficiente, a empresa desenvolveu a área de *Creatives Intelligence* ou *Creatives Intel*, que tem como principal função a otimização da performance dos criativos. Essa tarefa tem papel estratégico para a

empresa, mas ainda está em desenvolvimento e não possui processos totalmente otimizados.

3.2.1. Processo Atual

Atualmente, a atuação de *Creatives Intel* baseia-se na produção de diferentes iterações para cada criativo produzido e na testagem deles para geração de aprendizados replicáveis. As iterações dos vídeos são feitas uma a uma: altera-se apenas um fator por vez, como, por exemplo, apenas um personagem, ou apenas o cenário. Para aproximar ao máximo os resultados das campanhas de uma conclusão, ao se fazer uma iteração, procura-se manter o restante do criativo o mais próximo possível, alterando-se apenas um elemento. A Figura 4 exemplifica iterações realizadas para a definição do personagem em criativos do jogo *Sniper 3D*. Após cada fator ser testado separadamente, combina-se os vencedores em uma versão ótima do criativo, que é utilizada em campanhas.



Figura 4 - Iterações de personagem testadas em estático para o jogo *Sniper 3D*, da Wildlife Studios

3.2.1.1. *Produção de criativos*

A empresa possui seis células de produção de criativos: Estáticos, *Unity*, *Motion* 3D, Cinemático, *Motion* 2D e *Playables*. Abaixo, suas funções e principais características são resumidamente explicadas:

- Estáticos

Célula na qual são produzidos os estáticos, ou imagens sem movimento. O tempo de produção de estáticos é baixo e, por isso, utiliza-se essa equipe para realização de testes que podem ser feitos sem animação, como iterações de cores, de fontes para textos e de frases, além da produção de anúncios estáticos.

- *Unity*

Célula utilizada para produzir animações em 3D de baixa complexidade. *Unity* é o nome do *software* (*game engine*) que utiliza modelos e animações prontas, retiradas do próprio jogo ou compradas em *sites* especializados. Esse reaproveitamento é feito para ganhar velocidade na produção. Essa equipe produz vídeos utilizados em análises iniciais de hipóteses que podem, posteriormente, ser produzidas em uma célula que permite maior complexidade e maior qualidade de imagem e animação, gerando vídeos com maior potencial de sucesso.

- *Motion* 3D

Célula utilizada para produzir animações em 3D com complexidade média. O *software* utilizado por essa equipe permite a produção de conteúdo com menos apelo realista, sendo utilizado para criativos de jogos que se aproximam do estilo *cartoon*, como *Zooba* e *Castle Crush* (Figura 5) e para partes do vídeo com apelo mais gráfico, como exposição de promoções e possibilidades de *upgrade* (Figura 6). Possui um tempo de produção superior ao de *Unity* pois demanda a produção e animação de modelos, mas ainda atinge um tempo inferior às células de produções mais complexas.



Figura 5 – Exemplo do estilo cartoon de animação do jogo Castle Crush, da Wildlife Studios



Figura 6 - Exemplo de cena com apelo gráfico produzida pela célula de Motion 3D. Promoção de upgrade do jogo Tennis Clash, da Wildlife Studios

- Cinemático

Célula utilizada para produzir animações em 3D com muitos efeitos e grande complexidade. Possui um tempo de produção maior, mas produz vídeos com efeitos mais realistas e animações mais complexas, que geralmente possuem melhores resultados em campanha, principalmente para divulgação de jogos com estilo mais realista, como *Battle Tanks* e *Sniper 3D* (Figura 7).



Figura 7 - Exemplo de animação com apelo realista e grande complexidade do jogo Battle Tanks, da Wildlife Studios

- Motion 2D

Célula que permite a produção de animações em duas dimensões (2D, sem profundidade). Nela, geralmente são produzidos vídeos simples para jogos que não possuem visual gráfico tridimensional, como o *Math Learner* e o *Colorfy* (Figura 8) e, principalmente, edições e manipulações de vídeos prontos, como cortes, textos e adição de elementos bidimensionais (como os elementos da Figura 7 que estão acima do cenário e representam controles e informações do jogo; elementos gráficos exibidos na tela para transmitir informações ao jogador são chamados de HUD ou *Heads-Up Display*). A equipe de Motion 2D é muito utilizada para iterações feitas em vídeos de Motion 3D ou Cinemático.

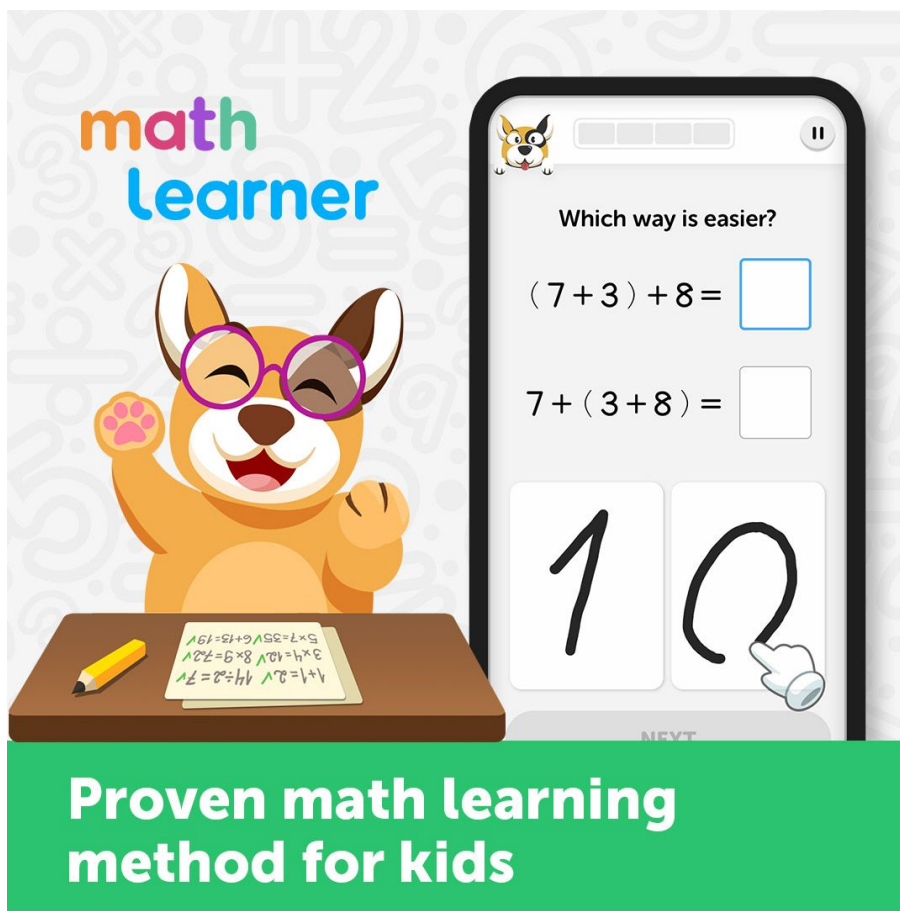


Figura 8 - Exemplo de animação bidimensional produzida em Motion 2D do jogo Math Learner, da Wildlife Studios

- *Playables*

Playables são os anúncios interativos ou jogáveis. Eles oferecem ao consumidor uma amostra da experiência do jogo. Essa célula é formada por artistas, responsáveis pela produção dos modelos utilizados nos anúncios, e engenheiros, responsáveis pela programação por meio da qual os elementos produzidos pelos artistas tornam-se interativos.

3.2.1.2. Hipóteses e Premissas

Os vídeos criados e suas iterações são sempre idealizados e produzidos pensando-se em hipóteses a serem provadas ou refutadas. Por ser um processo novo e pouco desenvolvido, as hipóteses testadas são sempre comparações binárias, com

apenas a alteração de um fator entre elas. Procura-se fazer a mesma alteração em diversos vídeos para que se tenha uma confiança maior nos resultados.

Em casos de alteração de fatores que consideram apenas a presença ou ausência de um elemento, como a presença ou não de dedos simulando um jogador no vídeo (Figura 9), a hipótese é que a presença da imagem do dedo do jogador torna o desempenho do criativo superior. Quando uma das versões apresenta resultados suficientemente superiores à outra, considera-se a hipótese confirmada ou rejeitada. Na empresa, considera-se uma versão suficientemente superior à outra quando ela apresenta um resultado 5% melhor que ela em testes. Esse número foi definido baseado no histórico de testes e em estudos realizados no passado.

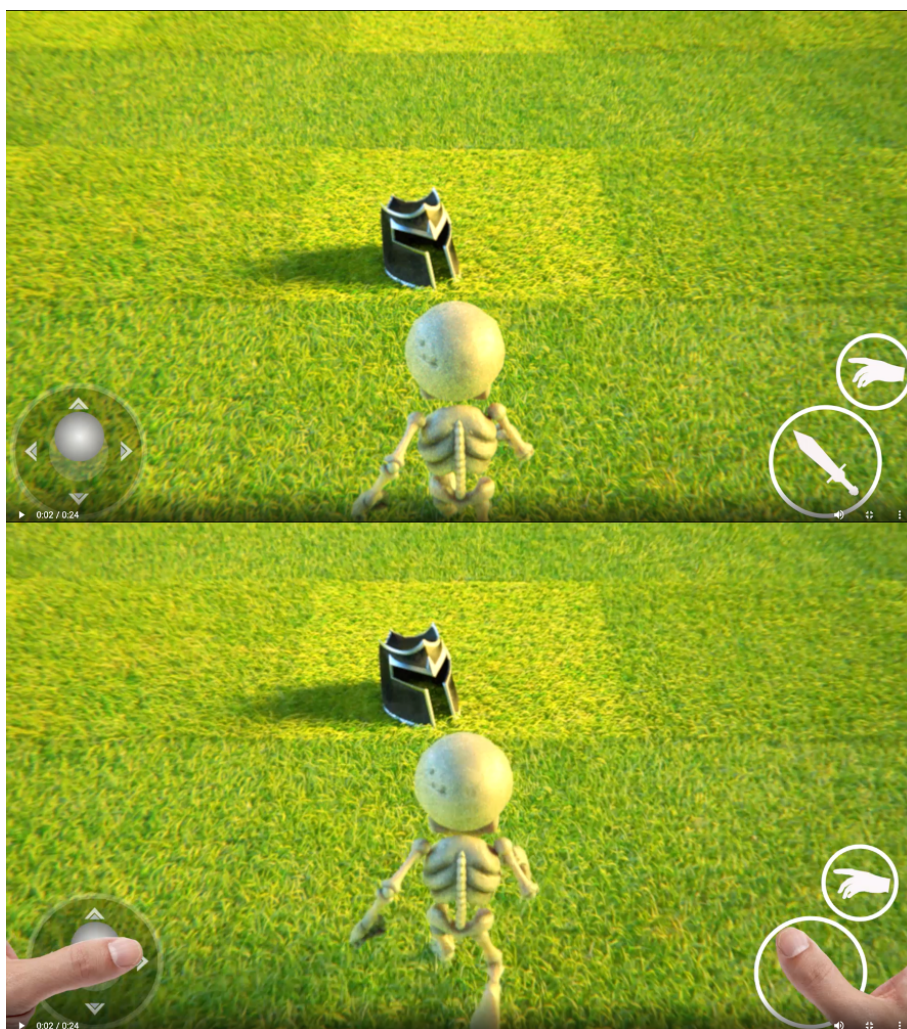


Figura 9 - Exemplo de alteração de fator binário em vídeo do jogo Castle Crush, da Wildlife Studios

Já para fatores que possuem múltiplas variações, como a alteração de uma personagem (como o exemplo da Figura 4), elabora-se diferentes hipóteses binárias

e estuda-se uma a uma separadamente. No exemplo da Figura 4, como existem quatro personagens, que serão chamados de Personagem A, Personagem B, Personagem C e Personagem D, as hipóteses testadas são:

- Personagem A é melhor que Personagem B;
- Personagem A é melhor que Personagem C;
- Personagem A é melhor que Personagem D;
- Personagem B é melhor que Personagem C;
- Personagem B é melhor que Personagem D;
- Personagem C é melhor que Personagem D.

Os resultados dessas comparações são analisados e define-se a partir deles a melhor personagem para a performance do anúncio.

Assume-se que a colinearidade entre fatores é desprezível nos testes de hipóteses binárias. Assim, cada uma delas é testada separadamente e se faz combinações entre as melhores, esperando-se que o resultado final seja positivo. Embora não seja verdade em todos os casos, essa premissa tem produzido resultados e, por isso, não foram despendidos recursos até hoje para testá-la e possivelmente alterar o processo utilizado.

3.2.1.3. Testes

Para se descobrir a eficiência de cada vídeo, são feitos os chamados “testes AB”. Cada anúncio presente em uma campanha é enviado para públicos estatisticamente semelhantes pelas redes de divulgação. Uma campanha é feita com, no mínimo, cinco anúncios, para verificar diversas hipóteses em um mesmo teste. As principais redes utilizadas para testes são o *Facebook* e a *Applovin*.

No Facebook, os testes possuem cinco vídeos cada, sendo um o controle (vídeo de melhor performance do jogo na rede). Um orçamento por anúncio é definido e a rede divulga cada criativo para públicos com perfis semelhantes até que o número de divulgações consuma o valor pré-estabelecido. Dessa forma, cada anúncio é enviado a um número de pessoas muito próximo e se comparam o número de *downloads*. A métrica utilizada para avaliar a performance dos vídeos é o CPI, ou seja,

a divisão do orçamento definido pela quantidade de pessoas que baixaram o jogo após assistir esse vídeo.

Já na Applovin, os testes são feitos com dez vídeos cada, sendo um o controle. Nessa rede, não se define o *budget* total, mas o próprio CPI, ou lance. Ou seja, define-se quanto se quer pagar por usuário que fizer o *download* do jogo. Com isso, a rede começa a divulgar os vídeos para públicos com perfis semelhantes, procurando manter uma quantidade próxima de impressões por criativo. A campanha deve ser pausada manualmente quando se atinge um volume de instalações suficiente e a métrica utilizada para analisar a qualidade dos criativos (uma vez que o CPI é definido e constante) é o IPM, ou seja, a proporção entre o número de pessoas que fez o *download* do jogo após assistir o anúncio e o número de pessoas que receberam o vídeo.

Os testes são organizados de forma a testar o maior número de hipóteses possível por vez, combinando-se os criativos de forma otimizada. Eles são feitos em países que possuem comportamento semelhante ao do público dos Estados Unidos, mas que possuem um custo de aquisição inferior ao americano. O custo necessário para se obter um número suficiente de impressões para que os resultados sejam estatisticamente confiáveis nesses países varia de trinta a cinquenta dólares por anúncio, o que, normalmente, possibilita uma média de trinta mil impressões por criativo. Dessa forma, no *Facebook*, um teste custa em média de US\$150 a US\$250 e atinge aproximadamente 150 mil pessoas. Já na *Applovin*, um teste custa de US\$300 a US\$500 e atinge cerca de 300 mil pessoas. Esses valores podem variar muito quando se altera os parâmetros definidos para o teste.

3.2.1.4. Análises

A comparação de hipóteses é o principal objetivo dos testes e o que guia os processos da área. Porém, diversas outras análises são feitas que possibilitam a otimização de performance dos criativos. Dentre elas está a análise do percentual de pessoas que assiste até determinado tempo do vídeo antes de fechá-lo, que possibilita o entendimento de que trechos possuem maior apelo, e testes feitos para públicos segmentados (dividindo-os por gênero, idade, país, gostos obtidos por meio do

comportamento deles em redes sociais e outras categorizações), que permite a avaliação e produção de anúncios focados em público específicos.

3.2.1.5. *Fraquezas do processo atual*

Atualmente, o processo de testagem de criativos traz resultados excelentes para a área e para a empresa. Consegue-se otimizar e melhorar constantemente a performance das campanhas e sabe-se, por meio de testes e comparações com bases de dados de criativos das redes, que os produzidos pela empresa estão entre os melhores do mundo. Entretanto, a área de *Creatives Intel* é muito recente e seus processos estão em constante estudo. Sabe-se que o que é feito hoje não é feito da maneira otimizada.

Primeiramente, por assumir-se que a correlação entre os fatores na performance de um criativo é desprezível, perde-se muito conhecimento e, possivelmente, os criativos mais eficientes nunca chegam a ser produzidos. Ao se testar cada hipótese separadamente, o número de vídeos produzidos e o volume de testes é muito superior ao mínimo necessário. Além disso, a realização de testes apenas para hipóteses binárias exclui uma grande gama de aprendizados que poderiam ser feitos estudando-se outras hipóteses.

A produção de vídeos que possuem combinações diferentes e complementares de diversas hipóteses possibilitaria uma melhoria expressiva nas duas primeiras fraquezas citadas. Seria possível entender e mensurar a correlação entre fatores testados e estudar diversas hipóteses com um número menor de vídeos produzidos e de testes realizados, possibilitando uma redução considerável de custo.

Para isso, foi aplicado um delineamento de experimentos, buscando-se atingir essas melhorias. Os resultados foram analisados e comparados com o modelo atual.

4. PLANEJAMENTO

Um experimento fatorial fracionado 2^{k-p} foi realizado e os resultados comparados com os obtidos por meio do processo atualmente utilizado pela empresa. Para isso, foi utilizado um vídeo de anúncio do jogo *Castle Crush* e oito fatores foram analisados em iterações do vídeo.

Castle Crush é um jogo de cartas e estratégia. Nele, o jogador conquista suas cartas por meio de batalhas e as acumula e coleciona. Cada carta corresponde a um personagem, que é adicionado ao exército do jogador. Cada partida é uma batalha entre dois jogadores em que cada um deve adicionar suas cartas estrategicamente ao tabuleiro para atacar o adversário e, ao mesmo tempo, proteger seu castelo. Cada carta possui características diferentes de ataque e defesa, e devem ser combinadas de diferentes formas para possibilitar a melhor estratégia de combate. A Figura 10 é uma captura de tela de uma partida do jogo, em que pode-se ver, nas laterais, os dois castelos, sendo o da esquerda do jogador principal e o da direita o do adversário. Na parte inferior estão as cartas do exército, que podem ser adicionadas em qualquer posição entre os castelos para atacar ou defender.



Figura 10 - Captura de tela do jogo *Castle Crush*, da Wildlife Studios

4.1. O VÍDEO BASE

O vídeo utilizado como base para as comparações é uma dramatização de uma estratégia de combate real do jogo. Esse tipo de vídeo é muito comum e costuma representar uma situação do jogo de forma dinâmica, ampliando a ação e tornando o vídeo mais atrativo. É comum que esses vídeos não representem de forma estritamente fiel as características do jogo, podendo ter variações que tornam o vídeo mais bonito ou eficiente. Na Wildlife Studios busca-se, porém, nunca se distanciar demais do jogo original para que o anúncio não seja considerado enganoso.



Figura 11 - Cartas dos personagens presentes no vídeo estudado: Warrior, Skeleton Swarm e Mud Elemental

No vídeo, o personagem principal é o guerreiro (*Warrior* – carta da esquerda na Figura 11), uma das cartas iniciais do jogo, e ele investe contra um castelo que está sendo defendido por um batalhão de esqueletos (*Skeleton Swarm* – carta central na Figura 11), uma das cartas mais populares. Ao se aproximar, o guerreiro esconde-se atrás de uma rocha e arremessa uma carta, indicando que um outro personagem foi ativado. Da carta arremessada surge o monstro de lama (*Mud Elemental* – carta da direita na Figura 11) que possui como principal característica uma elevada resistência a ataques. O guerreiro, então, utiliza dessa característica para aproximar-se do castelo sem ser atingido pelos ataques dos esqueletos, escondendo-se atrás do monstro de lama. Ao ver-se suficientemente próximo dos oponentes, ele retira-se de trás do parceiro e ataca o batalhão de esqueletos, derrotando-os e alcançando o castelo e vencendo a batalha.

A Figura 12 é um *storyboard* do vídeo base, que inclui os principais acontecimentos sequencialmente. Em ordem, as cenas mostradas são: guerreiro se aproximando do castelo do oponente, com o batalhão de caveiras em frente a ele; guerreiro escondendo-se atrás da rocha para se defender; guerreiro segurando a carta do monstro de lama; guerreiro arremessando a carta do monstro de lama; monstro de lama surgindo da carta arremessada; guerreiro posicionando-se atrás do monstro de lama para aproximar-se do castelo; guerreiro saindo de trás do monstro de lama para atacar os esqueletos; guerreiro atacando os esqueletos; e guerreiro atacando o castelo, após derrotar os esqueletos que o protegiam.



Figura 12 - Storyboard do vídeo estudado com principais acontecimentos

Ao final do vídeo, como de costume, existe o que é chamado de “cartela”, que é uma cena final que convida o usuário a fazer o *download* do jogo. A cartela aplicada ao vídeo estudado (Figura 13) foi a padrão do jogo, que contém um aparelho de telefone celular com uma animação de uma partida do jogo, a logomarca do jogo e um botão com a frase “*Download now*” (“Baixe agora”), que direciona para a loja de aplicativos quando clicado.

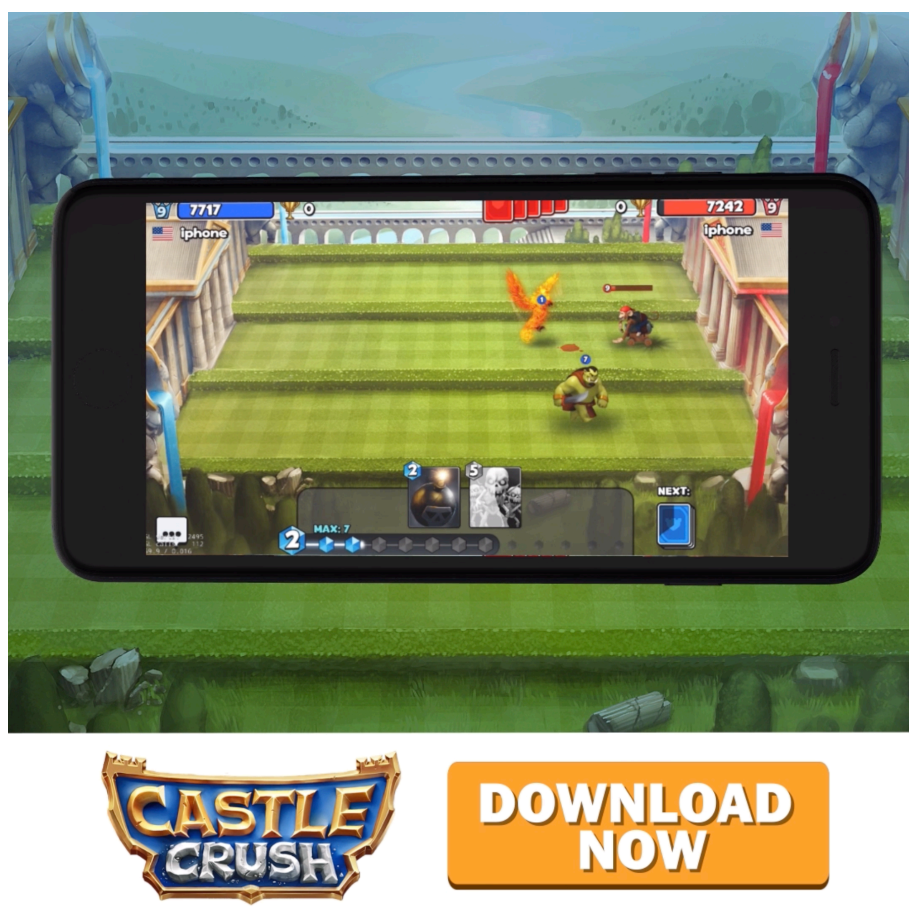


Figura 13 - Cartela aplicada ao vídeo base

O vídeo foi produzido na célula de Motion 3D, pois apresenta visual que se aproxima do estilo *cartoon* e complexidade média de animação. Sua cartela foi produzida na célula de Motion 2D, pois apresenta animações simples e elementos bidimensionais.

4.2. ITERAÇÕES

Oito fatores foram selecionados para serem testados. Eles são iterações feitas sobre o vídeo base e foram escolhidos com o objetivo de otimizar a performance do vídeo em campanhas. Os oito fatores são:

4.2.1. Alteração do posicionamento dos esqueletos

A primeira iteração proposta foi um reposicionamento dos esqueletos. Sabe-se, por meio de análises feitas previamente, que os esqueletos são extremamente populares no jogo e possuem um apelo visual que costuma melhorar a performance dos vídeos. Dessa forma, foi proposto que eles fossem posicionados sobre o castelo e não à frente dele, de forma a dá-los uma maior visibilidade. O contraste com o céu, por ter menos detalhes que o castelo, possibilitaria que eles ficassem mais visíveis, além de ser possível vê-los acima do monstro de lama quando ele surge (no vídeo base eles acabam escondidos pelo monstro) e, assim, espera-se que a performance do vídeo seja superior. Para que o enredo continue funcionando, nessa iteração, em vez de o cavaleiro contornar o monstro de lama, ele salta por cima dele e atinge os esqueletos no teto do castelo. A Figura 14 mostra cenas da versão do vídeo com a alteração do posicionamento dos esqueletos. Na primeira, a visibilidade dos esqueletos aumentada pelo contraste com o céu e por não serem encobertos pelo monstro de lama é facilmente visualizada. Na segunda, pode-se ver a alteração necessária no final do enredo, em que o cavaleiro salta sobre o monstro de lama para atacar os esqueletos.

Essa iteração foi produzida, assim como o vídeo, pela célula de Motion 3D, e possuiu um tempo de produção elevado, pois exige a reanimação de elementos, assim como adaptações de luz e sombras.



Figura 14 - Iteração de alteração do posicionamento dos esqueletos

4.2.2. Alteração do cenário

A segunda iteração proposta para análise é um teste de cenário. O jogo apresenta diversos cenários em que os castelos podem estar localizados. Os mais populares no jogo são o gramado visto no vídeo base e o cenário do inferno. Por isso, uma versão do vídeo no inferno também foi produzida para se avaliar qual dos dois cenários possui uma performance superior. A Figura 15 mostra a versão de cenário com temática de inferno.



Figura 15 - Iteração de alteração do cenário

Essa iteração também foi produzida inteiramente pela célula de Motion 3D. Seu tempo de produção concentrou-se no desenvolvimento do cenário e na adaptação de cores e luzes, pois as animações não foram alteradas.

4.2.3. Adição de *HUD*

A terceira iteração proposta é a adição de uma *HUD* ao vídeo. *HUD* ou *Heads-Up Display* é o nome dado a elementos gráficos exibidos na tela para transmitir informações ao jogador. Em outros testes já realizados na empresa, notou-se que quando se adiciona a *HUD* aos vídeos, transmite-se uma maior sensação de jogabilidade e costuma-se obter resultados superiores em campanhas. A *HUD*, em vídeos de dramatização de situações de jogo, como o estudado, pode afastar o criativo do jogo original. Nesses casos o resultado esperado é mais imprevisível, já foram vistos criativos que melhoram em performance e criativos que pioram pela adição de *HUD*. A Figura 16 mostra a *HUD* adicionada ao vídeo, destacada pelo contorno vermelho, na parte inferior da imagem. Nesse enredo, por se tratar de uma dramatização de situação do jogo, a *HUD* usada não pode ser a mesma do jogo, pois ela não se encaixaria, uma vez que o jogo possui um formato completamente diferente e não é jogado controlando-se um personagem individualmente, como é sugerido pelo vídeo. Por isso, foi criada uma *HUD* que fizesse sentido para esse vídeo especificamente.

Essa iteração foi produzida pela célula de Motion 2D, pois trata de elementos bidimensionais adicionados sobre o vídeo. Seu tempo de produção é baixo, pois sua complexidade é pequena.



Figura 16 - Iteração de adição de HUD

4.2.4. Adição de elementos gráficos de sinalização do jogador

A quarta iteração proposta segue o mesmo raciocínio da terceira (inclusão de *HUD*) e tem como objetivo aproximar o enredo de uma situação real de jogo. Para isso, foram adicionados elementos gráficos de sinalização do jogador muito comuns em *games*. Eles são uma forma circular posicionada sob o personagem que está sendo controlado e uma sinalização acima de sua cabeça, como ilustrado na Figura 17. Mais uma vez, sabe-se que esses elementos afastam o enredo da verossimilhança, pois transmitem a mensagem de que a dramatização feita no vídeo é o jogo real e isso pode ter um resultado negativo no desempenho das campanhas. Porém, existem diversos casos na empresa em que esse tipo de alteração possibilitou melhorias expressivas de performance e, portanto, seu teste é interessante para a área.



Figura 17 - Iteração de adição de elementos gráficos de sinalização do jogador

Assim como a iteração anterior, essa iteração foi produzida pela célula de Motion 2D e teve um tempo de produção baixo.

4.2.5. Inclusão de cena com destaque para a carta

A quinta iteração foi proposta pelos artistas que perceberam que, apesar de o guerreiro arremessar uma carta, o destaque para ela foi muito baixo. Por se tratar de um jogo em que se coleciona cartas, é interessante que haja um destaque maior para esse elemento. Para isso, foi produzida uma versão em que existe um corte na cena no momento em que o cavaleiro pega a carta. Ali, foi incluída uma cena curta com destaque para a carta, que é uma representação da visão do guerreiro naquele momento. Entende-se que esse corte pode quebrar o ritmo do anúncio, mas ao mesmo tempo pode dar maior destaque para a carta e otimizar o entendimento do jogo. A Figura 18 mostra a cena que foi incluída entre a terceira e a quarta cenas do *storyboard* mostrado na Figura 12.



Figura 18 - Iteração de adição de cena com destaque para a carta

Essa iteração foi produzida pela célula de Motion 3D, pois utiliza elementos tridimensionais. Sua animação, porém, é extremamente simples, e o tempo de produção concentrou-se na criação dos elementos.

4.2.6. Inclusão de cena de *upgrade* do cavaleiro

A sexta iteração proposta também segue um aprendizado feito em estudos anteriores que mostra que mostrar ao usuário a possibilidade de *upgrades* melhora a performance do jogo. Entende-se que isso ocorre porque dá ao jogador a sensação de possibilidade de progressão no jogo, que é algo muito buscado por jogadores. Para isso, foi criada uma cena de *upgrade* do personagem principal, o cavaleiro, que foi adicionada entre o final do enredo e a cartela (o que é chamado de pré-cartela). A cena mostra o cavaleiro recebendo armas melhores, um letreiro indicando seu nível aumentando e barras de estatísticas de habilidades, muito comuns em games, também sendo acrescentadas a cada vez que o botão “UPGRADE” é pressionado por

uma mão. A Figura 19 mostra os quatro níveis exibidos na cena de *upgrade* do cavaleiro.



Figura 19 - Iteração de inclusão de cena de *upgrade* do cavaleiro

Essa iteração também foi produzida pela célula de Motion 3D e apresentou um tempo de produção relativamente alto devido à complexidade da animação do personagem.

4.2.7. Inclusão de cena de *upgrade* do monstro de lama

A sétima iteração proposta segue o mesmo raciocínio e conceito da sexta (inclusão de cena de *upgrade* do cavaleiro), mas é feita com o monstro de lama. Como o monstro de lama não possui armas, os *upgrades* pelos quais ele passa são de evolução de personagem, em que ele se transforma em personagens ou combinações de personagens mais poderosos. A Figura 20 mostra os quatro níveis exibidos na cena de *upgrade* do monstro de lama.



Figura 20 - Iteração de inclusão de cena de *upgrade* do monstro de lama

Essa iteração, assim como a anterior, foi produzida pela célula de Motion 3D, mas apresentou um tempo de produção consideravelmente inferior pois teve animações mais simples.

4.2.8. Alteração da cartela

A oitava e última iteração proposta é a alteração da cartela do vídeo. Algumas cartelas diferentes foram testadas e a que alcançou melhores resultados nos testes foi selecionada para participar do experimento. A cartela selecionada é a mostrada na Figura 21.



Figura 21 - Iteração de alteração da cartela

Esta cartela foi produzida pela célula de Estáticos, que realizou a ilustração dos personagens, e aplicada ao vídeo pela célula de Motion 2D, que realizou uma animação simples dos elementos.

Para o delineamento de um experimento fatorial fracionado 2^{8-4} , faz-se o confundimento dos fatores a serem estudados com as combinações dos fatores fictícios utilizados no Quadro 4. Chamaremos os oito fatores selecionados pelas oito letras maiúsculas: A, B, C, D, E, F, G e H, fazendo as correspondências com as iterações no vídeo base de acordo com o Quadro 5.

Quadro 5 - Correspondência entre os fatores do modelo utilizado e as iterações propostas para o vídeo base

Fatores do experimento	Descrição da alteração
Fator A	Alteração do posicionamento dos esqueletos
Fator B	Alteração do cenário
Fator C	Adição de <i>HUD</i>
Fator D	Adição de elementos gráficos de sinalização do jogador
Fator E	Inclusão de cena com destaque para a carta
Fator F	Inclusão de cena de <i>upgrade</i> do cavaleiro
Fator G	Inclusão de cena de <i>upgrade</i> do monstro de lama
Fator H	Alteração de cartela

Para o experimento fatorial fracionado 2^{8-4} , quatro dos fatores estudados são atribuídos aos quatro fatores fictícios isolados: a, b, c e d. Os outros quatro são confundidos com as quatro interações entre três fatores: abc, abd, acd e bcd. Dessa forma, as 16 rodadas do experimento a ser realizado são mostradas no Quadro 6. Nela, os sinais indicam a presença ou não do fator em questão, como detalhado a seguir:

- para o fator A, o sinal negativo representa a versão do vídeo em que os esqueletos estão no chão e o sinal positivo a que eles estão sobre o castelo;
- para o fator B, o sinal negativo representa a versão do vídeo com o cenário do gramado e o sinal positivo o cenário de inferno;
- para o fator C, o sinal negativo representa a versão do vídeo sem a *HUD* e o sinal positivo a versão com a *HUD*;

- para o fator D, o sinal negativo representa a versão do vídeo sem os elementos gráficos de sinalização do jogador e o sinal positivo a versão com eles;
- para o fator E, o sinal negativo representa a versão do vídeo sem a cena com destaque para a carta e o sinal positivo a com a cena;
- para o fator F, o sinal negativo representa a versão do vídeo sem a cena de *upgrade* do cavaleiro e o sinal positivo a com a cena;
- para o fator G, o sinal negativo representa a versão do vídeo sem a cena de *upgrade* do monstro de lama e o sinal positivo a com a cena;
- para o fator H, o sinal negativo representa a versão do vídeo com a cartela original e o sinal positivo a com a cartela nova.

Quadro 6 - Sinais para as 16 rodadas do experimento delineado

Versões	a	b	c	d	abc	abd	acd	bcd
	A	B	C	D	E	F	G	H
1	-	-	-	-	-	-	-	-
2	+	-	-	-	+	+	+	-
3	-	+	-	-	+	+	-	+
4	+	+	-	-	-	-	+	+
5	-	-	+	-	+	-	+	+
6	+	-	+	-	-	+	-	+
7	-	+	+	-	-	+	+	-
8	+	+	+	-	+	-	-	-
9	-	-	-	+	-	+	+	+
10	+	-	-	+	+	-	-	+
11	-	+	-	+	+	-	+	-
12	+	+	-	+	-	+	-	-
13	-	-	+	+	+	+	-	-
14	+	-	+	+	-	-	+	-
15	-	+	+	+	-	-	-	+
16	+	+	+	+	+	+	+	+

As 16 versões de teste foram produzidas de acordo com o Quadro 6. Assim, a versão 1 corresponde ao vídeo base e cada uma das outras contém todas as iterações sugeridas em cada um dos fatores com sinal positivo.

Nesse experimento, os fatores individuais são confundidos com as interações de três fatores e as interações de dois fatores são confundidas entre si, quatro a quatro, de acordo com o Quadro 7, que detalha a correspondência entre os fatores fictícios do modelo com as interações confundidas. As interações de três ou mais fatores reais foram omitidas para simplificação da visualização, uma vez que são consideradas irrelevantes. Dessa forma, as conclusões a respeito dos fatores individuais são diretas. Já para as interações de dois fatores, pode-se identificar a influência delas em grupos de quatro e, caso alguma seja considerada relevante, pode-se estudá-la mais a fundo para entender qual ou quais das interações no confundimento são mais importantes.

Quadro 7 - Correspondência entre fatores fictícios do modelo e interações propostas confundidas

Fatores fictícios	Interações confundidas
a	A
b	B
c	C
d	D
ab	AB + CE + DF + GH
ac	AC + BE + DG + FH
ad	AD + BF + CG + EH
bc	AE + BC + DH + FG
bd	AF + BD + CH + EG
cd	AG + BH + CD + EF
abc	E
abd	F
acd	G
bcd	H
abcd	AH + BG + CF + DE

4.3.1. Blocagem

Os testes a serem realizados foram feitos na rede Applovin, que permite até 10 vídeos por teste. Como no projeto precisamos testar 16 versões, elas foram separadas em dois blocos de oito, de acordo com o modelo de blocagem sugerido por Box, Hunter e Hunter (2005). Nessa blocagem, os efeitos de ruído entre os dois blocos diferentes foram confundidos com os efeitos da interação de quatro fatores fictícios abcd que se confundem com as interações de dois fatores AH, BG, CF e DE. Dessa forma não foi possível avaliar o efeito dessas interações diretamente.

Para que os resultados do experimento sejam o mais aproveitáveis possível, a ordenação dos fatores foi feita pensando-se em quais interações entre dois deles são mais esperadas. Aos fatores A, B, C e D foram atribuídas alterações que se mantêm contínuas ao longo do vídeo todo para que se entenda a interação entre eles. Por outro lado, os fatores F e G, correspondem a iterações de pré-cartela, assim como o fator H, que também corresponde a uma alteração presente no trecho final do vídeo, estão juntos em um bloco pois espera-se que interfiram mais nos resultados uns dos outros. Dessa forma, os dois blocos agrupam alterações feitas em momentos distintos dos vídeos e, portanto, espera-se que os resultados confundidos com o efeito da blocagem sejam pouco relevantes.

O Quadro 8 repete os sinais indicados para cada uma das versões mostradas no Quadro 7, incluindo, também, os sinais para a interação dos quatro fatores fictícios abcd, que se confunde com as interações de dois fatores reais AH, BG, CF e DE. Esses sinais correspondem à multiplicação dos sinais dos fatores a (A), b (B), c (C) e d (D) para cada uma das versões. Para os casos em que esse sinal é positivo, a versão foi alocada no bloco I, enquanto nos casos em que o sinal é negativo, a versão foi alocada no bloco II.

5. EXECUÇÃO E RESULTADOS

Para a realização da comparação de processos, foram feitos dois testes: o primeiro segue o modelo atual adotado na empresa e o segundo o experimento sugerido.

Para o primeiro teste, chamado de “Modelo Atual”, oito vídeos foram produzidos, cada um com apenas uma das oito iterações sugeridas, como ilustrado pelo Quadro 9. Cada uma das oito versões foi testada junto com o vídeo base e, ao final, uma versão foi produzida com todos os fatores que obtiveram um resultado positivo no teste (Versão Final do Modelo Atual).

Quadro 9 - Versões produzidas para o Modelo Atual

Versões	A	B	C	D	E	F	G	H
1	+	-	-	-	-	-	-	-
2	-	+	-	-	-	-	-	-
3	-	-	+	-	-	-	-	-
4	-	-	-	+	-	-	-	-
5	-	-	-	-	+	-	-	-
6	-	-	-	-	-	+	-	-
7	-	-	-	-	-	-	+	-
8	-	-	-	-	-	-	-	+

Já para o teste feito no modelo proposto, chamado de “Experimento”, as versões foram produzidas de acordo com o Quadro 6, totalizando 15 novas versões, além do vídeo base (versão 1 no quadro). Elas foram testadas em dois blocos de 8 versões e os resultados foram comparados para se produzir um vídeo com performance otimizada (Vídeo Final do Experimento). Nesse teste, além dos resultados individuais de cada fator, as interações entre fatores foram levadas em consideração para a produção do vídeo final.

Ao final da comparação, os vídeos finais foram testados junto com o vídeo base e a performance deles foi comparada. A análise dos processos levará em conta, além desse resultado, o custo de produção das versões e o custo dos testes necessários.

Dessa forma, foi avaliado qual dos dois processos (atual ou sugerido) é preferível para a empresa adotar.

5.1. TESTAGEM

Uma vez que os vídeos foram produzidos, teve início a fase de testes. Como explicado anteriormente, a *adnet* escolhida para a testagem foi a *Applovin*, que permite até dez criativos por campanha. Dessa forma, os oito criativos do Modelo Atual puderam ser testados juntos e para o Experimento, os 16 vídeos produzidos foram testados em dois blocos de 8. Sabe-se que é comum existir diferenças entre testes, causadas por variações como data de realização e público alcançado. Porém, baseado em estudos já realizados pela empresa, é tomado como premissa que a proporção de instalações entre os vídeos em cada teste mantém-se aproximadamente constante, o que é normalmente confirmado por testes realizados. O *IPM* (*Impressions per Mile* - instalações obtidas por 1000 impressões), resultado obtido como variável resposta dos testes em *Applovin*, pode, portanto, ser comparado entre testes diferentes, pois trata-se de uma proporção que se espera que se mantenha aproximadamente constante para cada anúncio independentemente dos testes, apesar de os valores das impressões e das instalações poderem variar consideravelmente entre eles.

Com o objetivo de reduzir os possíveis ruídos e obter-se resultados mais precisos, os testes para cada modelo foram repetidos duas vezes, em campanhas diferentes. Buscou-se obter um volume de impressões próximo de trinta mil, que é o valor buscado no modelo tradicional de testagem, e os resultados em *IPM* foram comparados com os do vídeo base, para que se possa avaliar o quanto cada fator afetou o *IPM*.

5.1.1. Testes e resultados para o Modelo Atual

Foram realizados dois testes de acordo com o padrão corrente de testagem para os oito vídeos produzidos. Neles, o vídeo base, sem nenhuma iteração, foi adicionado como controle para se obter a variação de *IPM* causada pelos fatores. Os

Quadros Quadro 10 e Quadro 11 mostram o volume de impressões e instalações obtidas para cada um dos oito vídeos e para o vídeo base (chamado de versão 0) nos dois testes realizados para o Modelo Atual (chamados de Teste 1 e Teste 2).

Quadro 10 - Resultados do Teste 1 - Modelo Atual

Versões	A	B	C	D	E	F	G	H	Impressões	Instalações
0	-	-	-	-	-	-	-	-	29435	172
1	+	-	-	-	-	-	-	-	29456	188
2	-	+	-	-	-	-	-	-	29979	154
3	-	-	+	-	-	-	-	-	31189	190
4	-	-	-	+	-	-	-	-	28551	178
5	-	-	-	-	+	-	-	-	29353	177
6	-	-	-	-	-	+	-	-	28961	175
7	-	-	-	-	-	-	+	-	31167	186
8	-	-	-	-	-	-	-	+	30677	188

Quadro 11 - Resultados do Teste 2 - Modelo Atual

Versões	A	B	C	D	E	F	G	H	Impressões	Instalações
0	-	-	-	-	-	-	-	-	33176	199
1	+	-	-	-	-	-	-	-	33037	213
2	-	+	-	-	-	-	-	-	31283	183
3	-	-	+	-	-	-	-	-	30978	222
4	-	-	-	+	-	-	-	-	32770	195
5	-	-	-	-	+	-	-	-	32327	208
6	-	-	-	-	-	+	-	-	32920	215
7	-	-	-	-	-	-	+	-	32259	213
8	-	-	-	-	-	-	-	+	30699	211

A partir disso, foram calculados o *IPM* médio dos testes e a razão do *IPM* total em relação ao controle (versão 0) para cada um dos vídeos, conhecida como *IPM ratio*. O Quadro 12 mostra os valores obtidos para cada versão e a Figura 22 ilustra a variação de *IPM* entre as versões testadas.

Quadro 12 - IPM total e IPM ratio para os vídeos do Modelo Atual

Versões	A	B	C	D	E	F	G	H	IPM médio	IPM ratio
0	-	-	-	-	-	-	-	-	5.925	1.000
1	+	-	-	-	-	-	-	-	6.417	1.083
2	-	+	-	-	-	-	-	-	5.501	0.928
3	-	-	+	-	-	-	-	-	6.627	1.118
4	-	-	-	+	-	-	-	-	6.083	1.027
5	-	-	-	-	+	-	-	-	6.242	1.053
6	-	-	-	-	-	+	-	-	6.302	1.064
7	-	-	-	-	-	-	+	-	6.291	1.062
8	-	-	-	-	-	-	-	+	6.501	1.097

IPM ratio por Versão - Modelo Atual

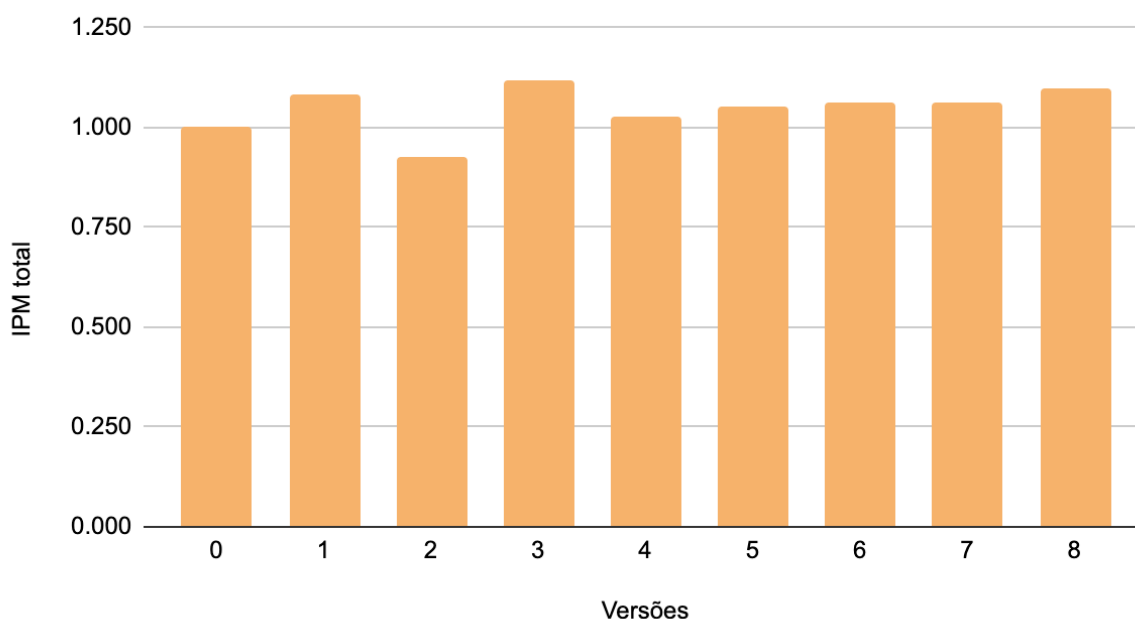


Figura 22 - IPM ratio por versão - Modelo Atual

Os resultados obtidos pelos dois testes determinam que apenas o fator 2 não otimiza os resultados do vídeo, que o fator 3 é o que mais melhora a performance, seguido pelo 8 e, depois, pelo 1. Os fatores 4, 5, 6 e 7 possuem um impacto positivo, mas mais baixos que os últimos 3 citados.

Quadro 15 - Resultados do Teste 5 - Experimento, Bloco II

Versões	A	B	C	D	E	F	G	H	Impressões	Instalações
2	+	-	-	-	+	+	+	-	27313	195
3	-	+	-	-	+	+	-	+	27359	180
5	-	-	+	-	+	-	+	+	27483	213
8	+	+	+	-	+	-	-	-	27907	188
9	-	-	-	+	-	+	+	+	28855	200
12	+	+	-	+	-	+	-	-	26171	169
14	+	-	+	+	-	-	+	-	27018	202
15	-	+	+	+	-	-	-	+	27387	173

Quadro 16 - Resultados do Teste 6 - Experimento, Bloco II

Versões	A	B	C	D	E	F	G	H	Impressões	Instalações
2	+	-	-	-	+	+	+	-	28544	206
3	-	+	-	-	+	+	-	+	29210	190
5	-	-	+	-	+	-	+	+	28625	218
8	+	+	+	-	+	-	-	-	30724	205
9	-	-	-	+	-	+	+	+	29582	208
12	+	+	-	+	-	+	-	-	28382	184
14	+	-	+	+	-	-	+	-	28955	221
15	-	+	+	+	-	-	-	+	30050	194

A Figura 23 ilustra a variação de *IPM* entre as versões testadas no Experimento.

IPM ratio por Versão - Experimento

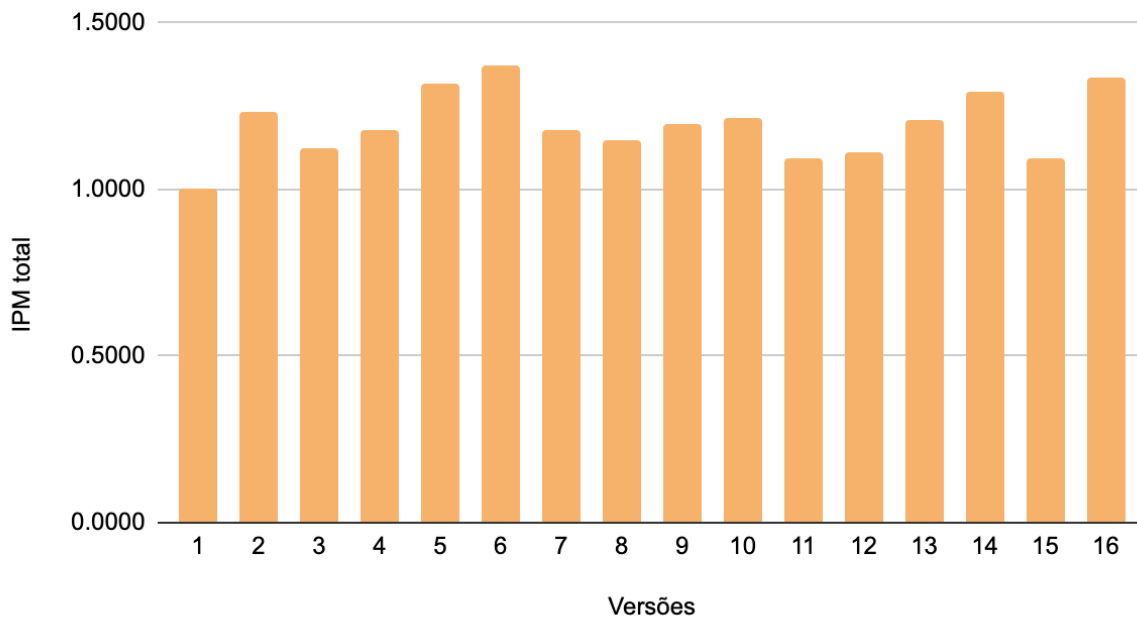


Figura 23 – *IPM ratio por versão - Experimento*

Por se tratar de um experimento delineado, nenhuma conclusão direta sobre fatores individuais pode ser tirada da análise visual do gráfico com os *IPMs* obtidos, pois nenhuma das versões possui apenas uma alteração feita. Pode-se, porém, constatar que o vídeo 1 do Experimento, que é o vídeo de controle, sem alterações, obteve um *IPM* próximo com o seu equivalente no teste do Modelo Atual (o vídeo 0, com as mesmas características), o que reforça o fato de que a influência de dia de aplicação e público atingido não leva a grandes variações de *IPM*, apesar das diferenças entre impressões e instalações em testes diferentes.

6. ANÁLISE

Com os seis testes principais realizados, cada modelo será analisado de acordo com a sua proposta e os dois processos serão comparados.

6.1. ANÁLISE DOS RESULTADOS DO MODELO ATUAL

O Modelo Atual será tratado conforme o processo corrente na empresa. Para isso, compara-se a variação de *IPM* de cada versão em relação ao vídeo controle (versão 0), como ilustrado na Figura 24.

Variação de IPM em relação ao controle (vídeo 0) - Modelo Atual



Figura 24 - Variação de IPM em relação ao controle (vídeo 0) - Modelo Atual

A análise usualmente feita leva em consideração quais fatores obtiveram uma variação de pelo menos 5% em relação ao controle. Assim, a versão 4, com adição de elementos gráficos de sinalização do jogador, que obteve uma variação positiva de

2,66% em relação ao vídeo original é considerada um empate com o controle. Entende-se, portanto, que essa alteração não tem efeito relevante sobre a performance do vídeo e, portanto, ela é descartada. Já a versão 2, com o cenário com a temática do inferno, obteve uma variação negativa de 7,16% em relação ao controle. Ela é considerada uma derrota em relação ao controle e, portanto, a hipótese do cenário de inferno é rejeitada. As demais alterações tiveram um efeito positivo superior a 5% na performance do vídeo. Por isso, todas elas são consideradas vitoriosas e suas hipóteses são confirmadas.

O processo demanda que, nesse momento, seja produzida uma versão com todas as alterações vencedoras encontradas pelos testes. Portanto, foi produzido um vídeo com as alterações A, C, E, F, G e H. Além disso, baseado em conhecimentos já adquiridos pela área, foram pedidas mais duas versões: uma com os fatores A, C, E, F e H alterados e outra com apenas os fatores A, C e H.

A primeira foi adicionada pois sabe-se que os fatores F e G são semelhantes: ambos são “pré-cartelas” de *upgrade*, em que se altera apenas o personagem. Nesses casos, sabe-se que normalmente adicionar-se apenas uma delas causa um impacto superior à inclusão de duas pré-cartelas no mesmo vídeo. Esse conhecimento não é comprovado teoricamente, mas foi desenvolvido organicamente pela empresa ao longo do tempo. Já a segunda versão extra pedida exclui as alterações que tiveram um efeito muito próximo de 5% de variação do *IPM* em relação ao vídeo original. Esse tipo de versão é pedido para se avaliar se a aplicação de todos os fatores testados produz resultado, pois devido ao aumento no tempo de produção, prefere-se versões que possuam menos iterações em relação ao vídeo base. Além disso, a segurança de que os resultados muito próximos de 5% são realmente relevantes é baixa.

Assim, três versões resultantes da análise do Modelo Atual foram pedidas. Elas foram testadas com as versões resultantes da análise do Experimento para comparação de resultados obtidos. Todas as versões pedidas estão detalhadas no Quadro 23, no item 6.2.5 (Resultados dos testes finais).

6.2. ANÁLISE DOS RESULTADOS DO EXPERIMENTO

A análise do Experimento seguiu Box, Hunter e Hunter (2005). A primeira etapa é o cálculo do efeito de cada fator ou interação de fatores a partir dos resultados obtidos. Esse valor é calculado como a diferença da média dos resultados obtidos nas versões em que cada fator ou interação tem o sinal positivo (está presente) e a média dos resultados obtidos nas versões em que ele é negativo (ausente), como explicado anteriormente. Isso é facilmente feito a partir do quadro de sinais proposto.

6.2.1. Efeitos de Fatores Individuais

Primeiro, são calculados os efeitos dos fatores individuais, isso é, o impacto que cada uma das alterações isoladas causa no resultado do vídeo. Como explicado anteriormente, para se calcular isso, deve-se subtrair a média dos resultados obtidos com o sinal negativo para cada fator da média dos resultados obtidos com o sinal positivo. Assim, utilizando-se os *IPM ratios* encontrados no Quadro 17, obtém-se o efeito de A: $l_A = 7,73\%$, que significa que o posicionamento dos esqueletos sobre o castelo possui um efeito de melhoria de performance de 7,73% em média.

Analogamente, os demais efeitos individuais foram calculados e obteve-se os dados exibidos no Quadro 18.

Quadro 18 - Efeitos individuais calculados para o Experimento

Efeito	Valor
l_A	7,73%
l_B	- 6,34%
l_C	10,94%
l_D	0,08%
l_E	3,16%
l_F	5,87%
l_G	8,06%
l_H	6,35%

6.2.2. Efeitos de Interações de Dois Fatores

Os efeitos das interações de dois fatores são calculados da mesma forma que para fatores individuais, utilizando-se os sinais do Quadro 4 e os confundimentos propostos no Quadro 7, obtém-se, assim, os valores exibidos no Quadro 19.

Quadro 19 - Efeitos de interações de dois fatores calculados para o Experimento

Efeitos	Valor
$l_{AB}, l_{CE}, l_{DF}, l_{GH}$	- 2,07%
$l_{AC}, l_{BE}, l_{DG}, l_{FH}$	0,79%
$l_{AD}, l_{BF}, l_{CG}, l_{EH}$	- 0,37%
$l_{AE}, l_{BC}, l_{DH}, l_{FG}$	- 4,73%
$l_{AF}, l_{BD}, l_{CH}, l_{EG}$	0,52%
$l_{AG}, l_{BH}, l_{CD}, l_{EF}$	- 2,82%
$l_{AH}, l_{BG}, l_{CF}, l_{DE}$	0,96%

6.2.3. Variância e Desvio Padrão

O passo seguinte é o cálculo do erro padrão dos resultados entre os testes diferentes feitos com cada amostra. Ele tem como objetivo definir qual intervalo dos efeitos deve ser considerado relevante na performance dos vídeos. É análogo ao valor de 5% utilizado no Modelo Atual, que foi definido pela empresa baseado no histórico, mas é específico para esse experimento.

Para isso, como explicado anteriormente, calcula-se primeiro a variância para os resultados de *IPM* para cada par de testes ao qual cada versão foi submetida. Para os criativos do bloco I, a primeira rodada corresponde ao teste 3 e a segunda ao teste 4. Para o bloco II, a primeira rodada corresponde ao teste 5 e a segunda ao teste 6. O Quadro 20 mostra os resultados calculados.

Quadro 20 - *IPM por rodada e variâncias*

Versões	A	B	C	D	E	F	G	H	<i>IPM</i> – 1ª rodada	<i>IPM</i> – 2ª rodada	σ_{IPM}^2
1	-	-	-	-	-	-	-	-	5,8278	5,5763	0,031633
2	+	-	-	-	+	+	+	-	7,1395	7,2169	0,003001
3	-	+	-	-	+	+	-	+	6,5792	6,5046	0,002780
4	+	+	-	-	-	-	+	+	6,9071	6,7533	0,011815
5	-	-	+	-	+	-	+	+	7,7502	7,6157	0,009048
6	+	-	+	-	-	+	-	+	7,9915	8,1385	0,010805
7	-	+	+	-	-	+	+	-	6,8876	7,0975	0,022047
8	+	+	+	-	+	-	-	-	6,7367	6,6723	0,002071
9	-	-	-	+	-	+	+	+	6,9312	7,0313	0,005010
10	+	-	-	+	+	-	-	+	7,0901	6,6982	0,076795
11	-	+	-	+	+	-	+	-	6,3588	6,8397	0,115664
12	+	+	-	+	-	+	-	-	6,4575	6,4830	0,000324
13	-	-	+	+	+	+	-	-	6,9446	7,2614	0,050157
14	+	-	+	+	-	-	+	-	7,4765	7,6325	0,012174
15	-	+	+	+	-	-	-	+	6,3169	6,4559	0,009666
16	+	+	+	+	+	+	+	+	7,8395	7,6848	0,011953

Como se trata da variância dos *IPMs* e sabe-se que eles se mantêm aproximadamente constantes para cada criativo, percebe-se que os valores encontrados são relativamente baixos. Além disso, por se tratar de apenas dois testes para cada versão, os valores das variâncias podem sofrer alteração considerável.

Como a métrica utilizada para se obter os resultados finais do experimento foi o *IPM ratio*, estimou-se, a partir dos *IPMs* encontrados para cada teste e suas variâncias, a variância do *IPM ratio* por meio da aproximação da variância de uma razão proposta por Seltman (2018), abaixo:

$$\sigma_{A/B}^2 = \frac{A^2}{B^2} \left[\frac{\sigma_A^2}{A^2} + \frac{\sigma_B^2}{B^2} - 2 \frac{cov(A, B)}{AB} \right]$$

Como os *IPM ratios* calculados são todos em relação ao *IPM* da versão 1, B e σ_B^2 não se alteram, sendo os valores relacionados à primeira versão, enquanto A e σ_A^2 correspondem a cada uma das outras versões. Assim, as variâncias calculadas para cada *IPM ratio* estão indicados no Quadro 21.

Quadro 21 - Variâncias dos *IPM ratios* para cada versão

Versões	A	B	C	D	E	F	G	H	<i>IPM ratio</i>	$\sigma_{IPM\ ratio}^2$
1	-	-	-	-	-	-	-	-	1,0000	0.0009672
2	+	-	-	-	+	+	+	-	1,2553	0.0019897
3	-	+	-	-	+	+	-	+	1,1437	0.0010221
4	+	+	-	-	-	-	+	+	1,1960	0.0010378
5	-	-	+	-	+	-	+	+	1,3432	0.0013267
6	+	-	+	-	-	+	-	+	1,4088	0.0030464
7	-	+	+	-	-	+	+	-	1,2212	0.0031023
8	+	+	+	-	+	-	-	-	1,1720	0.0011019
9	-	-	-	+	-	+	+	+	1,2208	0.0020645
10	+	-	-	+	+	-	-	+	1,2092	0.0019399
11	-	+	-	+	+	-	+	-	1,1495	0.0069403
12	+	+	-	+	-	+	-	-	1,1314	0.0013588
13	-	-	+	+	+	+	-	-	1,2386	0.0045255
14	+	-	+	+	-	-	+	-	1,3214	0.0028538
15	-	+	+	+	-	-	-	+	1,1173	0.0021001
16	+	+	+	+	+	+	+	+	1,3586	0.0013430

Calculadas as variâncias para cada fator do experimento, pode-se estimar a variância de cada efeito calculado como sugerido por Box, Hunter e Hunter (2005). Sendo s^2 a média simples das variâncias dos *IPM ratios* encontradas, temos que, como cada efeito encontrado ($l_A, l_B, l_C \dots$) é a diferença entre duas médias de oito observações cada, a variância estimada de cada efeito é:

$$\sigma_{efeito}^2 = \left(\frac{1}{8} + \frac{1}{8}\right) s^2 = \left(\frac{0,002295}{4}\right) = 0,000574$$

A partir da variância de cada efeito, encontra-se o erro padrão a partir de sua raiz quadrada. Aqui temos que o erro padrão dos efeitos calculados é igual a $\sqrt{0,000574} = 0,023953$. E, assim, temos que os efeitos calculados com erros são os mostrados no Quadro 22.

Quadro 22 - Efeitos e erros finais calculados para o Experimento

Efeitos	Valor
l_A	$7,73\% \pm 2,40\%$
l_B	$-6,34\% \pm 2,40\%$
l_C	$10,94\% \pm 2,40\%$
l_D	$0,08\% \pm 2,40\%$
l_E	$3,16\% \pm 2,40\%$
l_F	$5,87\% \pm 2,40\%$
l_G	$8,06\% \pm 2,40\%$
l_H	$6,35\% \pm 2,40\%$
$l_{AB}, l_{CE}, l_{DF}, l_{GH}$	$-2,07\% \pm 2,40\%$
$l_{AC}, l_{BE}, l_{DG}, l_{FH}$	$0,79\% \pm 2,40\%$
$l_{AD}, l_{BF}, l_{CG}, l_{EH}$	$-0,37\% \pm 2,40\%$
$l_{AE}, l_{BC}, l_{DH}, l_{FG}$	$-4,73\% \pm 2,40\%$
$l_{AF}, l_{BD}, l_{CH}, l_{EG}$	$0,52\% \pm 2,40\%$
$l_{AG}, l_{BH}, l_{CD}, l_{EF}$	$-2,82\% \pm 2,40\%$
$l_{AH}, l_{BG}, l_{CF}, l_{DE}$	$0,96\% \pm 2,40\%$

6.2.4. Análise dos resultados obtidos para o Experimento

Com os resultados obtidos, pode-se analisar cada efeito a fim de compreender as melhores combinações para um vídeo com performance otimizada. Como as alterações propostas para o vídeo foram baseadas em estudos e conhecimentos prévios, é esperado que a maior parte delas possua um impacto positivo sobre os resultados do anúncio.

Com os erros calculados e iguais a $\pm 2,40\%$, considera-se relevantes para a performance apenas os efeitos que possuem valores superiores a $2,40\%$ ou inferiores a $-2,40\%$. Assim, o único efeito individual considerado desprezível e, portanto, abandonado é o D (adição de efeitos gráficos para sinalização do jogador). Já o efeito B (alteração de cenário para versão do inferno) foi o único efeito que possuiu impacto negativo na performance do vídeo e, portanto, assim como o efeito D, foi abandonado. Todos os outros efeitos individuais obtiveram resultados positivos e seriam, portanto, adicionados a uma versão otimizada, que conteria os fatores A, C, E, F, G e H.

Dos efeitos individuais, o que se acredita possuir maior impacto positivo é o C (adição de *HUD*), seguido pelo G (inclusão de cena de *upgrade* do monstro de lama), e, depois, pelo A (alteração de posicionamento dos esqueletos). O que possui menor impacto positivo, baseado nos resultados obtidos é o fator E (inclusão de cena com destaque para carta), seguido pelo F (inclusão de cena de *upgrade* do cavaleiro) e, depois, pelo H (alteração da cartela).

A análise segue com os efeitos de interações entre dois fatores e vê-se que a maioria (20 de 28) é considerada irrelevante, ou seja, possui o resultado dentro da margem de erro do experimento. Isso reforça a hipótese de que, quanto maior a quantidade de fatores na interação, menor o impacto dela no resultado. Além disso, as quatro interações confundidas com os efeitos da blocagem ($l_{AH} = l_{BG} = l_{CF} = l_{DE} = 0,96\% \pm 2,40\%$) encontram-se no intervalo considerado sem impacto. Isso reforça a hipótese de que, apesar das diferenças entre testes, os valores de *IPM* para cada criativo mantêm-se aproximadamente constante.

As interações consideradas relevantes são:

$$\begin{aligned} l_{AE} = l_{BC} = l_{DH} = l_{FG} &= -4,73\% \pm 2,40\% \\ l_{AG} = l_{BH} = l_{CD} = l_{EF} &= -2,82\% \pm 2,40\% \end{aligned}$$

O ideal, nesse momento, seria uma análise mais profunda para entender como cada uma das interações confundidas afeta a performance e contribui para esse efeito negativo. Isso foi sugerido para a empresa e poderá ser feito nos próximos meses. Para essa monografia, porém, devido a limitações de orçamento, somente foram formuladas hipóteses sobre o motivo do efeito negativo.

Para o primeiro conjunto de confundimentos, que teve maior impacto ($l_{AE} = l_{BC} = l_{DH} = l_{FG}$), tem-se que, caso as interações que incluam os fatores B e D sejam

as mais relevantes, nenhuma ação é necessária, pois esses fatores foram abandonados. Delas, a interação entre B e C (alteração de cenário e adição de *HUD*) é considerada a com maior possibilidade de ter impacto negativo, dado que a alteração do cenário muda o esquema de cores e iluminação do vídeo todo e, como a *HUD* é um elemento que aparece por cima dele, pode-se perder contraste e visibilidade, afetando-se negativamente seu impacto. Além disso, a interação entre D e H (inclusão de elementos gráficos de sinalização do jogador e alteração da cartela) aparecem em momentos distintos do vídeo, trata-se de elementos completamente independentes. Portanto, espera-se que ela seja mais irrelevante. Em ambos os casos, dado que um dos elementos da interação não estará presente na versão otimizada do anúncio, não se espera que exista efeito negativo.

Já para as duas interações que não incluem fatores abandonados, temos que aquela entre os fatores A e E (alteração do posicionamento dos esqueletos e inclusão de cena com destaque para a carta) encaixa-se no mesmo caso da interação entre D e H. São fatores que não aparecem simultaneamente e tratam de elementos independentes do vídeo, e portanto, espera-se que possuam pouca interação. Já aquela entre F e G é exatamente o caso explicado na análise do Modelo Atual, em que se acredita que vídeos com duas pré-cartelas não possuem resultados bons pois elas competem entre si e podem confundir o usuário. Portanto, a hipótese é de que a interação entre F e G é a mais relevante.

Com essas análises, além da versão com todos os fatores com impacto positivo (A, C, E, F, G e H), outras três foram produzidas. A primeira excluindo a alteração E (inclusão de cena com destaque para a carta) que corresponde ao fator com menor efeito positivo da interação entre A e E. A segunda excluindo a alteração F (cena de *upgrade* do cavaleiro), fator com menor efeito positivo da interação entre F e G. E a terceira excluindo ambos os fatores E e F.

Para o segundo conjunto de confundimentos com impacto negativo ($l_{AG} = l_{BH} = l_{CD} = l_{EF}$), uma análise análoga foi realizada. Das interações que possuem fatores abandonados (B e D), a entre as alterações B e H (alteração de cenário e alteração da cartela) encontra-se no caso de elementos independentes e que não aparecem simultaneamente no vídeo e, portanto, espera-se que possua pouco impacto. Já a interação entre C e D (inclusão de *HUD* e adição de elementos gráficos de sinalização do jogador) trata de elementos que aparecem simultaneamente durante a maior parte do vídeo. Ambos correspondem a elementos gráficos

adicionados sobre o enredo principal e entende-se que, quando ambos estão presentes, pode-se obter uma imagem mais poluída, que distrai o usuário do enredo principal. Portanto, espera-se que essa seja a interação com maior impacto e, portanto, nenhuma alteração seria necessária.

As interações que não incluem fatores abandonados são entre A e G (alteração de posicionamento dos esqueletos e inclusão de cena de *upgrade* do monstro de lama) e entre E e F (inclusão de cena com destaque para a carta e inclusão de cena de *upgrade* do cavaleiro). Ambas se encaixam no caso de elementos independentes e que não aparecem simultaneamente e, portanto, espera-se que possuam pouco impacto na performance final.

A partir desses resultados, três novas versões foram pedidas. A primeira sem a alteração A (alteração do posicionamento dos esqueletos), que possui menor impacto entre as duas da interação entre A e G. A segunda sem o fator E (inclusão de cena com destaque para a carta), que possui menor impacto entre as duas alterações da interação entre E e F. E a terceira sem ambos os fatores A e E.

Como a versão em que se exclui o fator E se repete nas análises dos dois conjuntos de interações, as versões extras pedidas totalizam cinco e encontram-se detalhadas no Quadro 23, no item 6.2.5 (Resultados dos testes finais).

6.2.5. Resultados dos testes finais

A partir das análises dos resultados do experimento, portanto, nove versões do vídeo foram produzidas. No Quadro 23 estão as versões pedidas para a segunda rodada de testes. Nela, a versão I é a gerada a partir da análise do Modelo Atual e apenas dos fatores individuais do Experimento. As versões II e III são as pedidas a partir de conhecimentos prévios aplicados sobre os resultados do Modelo Atual. A versão IV corresponde à versão em que se elimina o fator E, pedida a partir da análise das duas interações de fatores com maior efeito. As versões V e VI são as outras duas pedidas para análise do conjunto de interações com maior impacto e as VII e VIII são as pedidas para análise do segundo conjunto de interações com maior impacto.

Quadro 23 - Versões produzidas para os testes finais

Versões	A	B	C	D	E	F	G	H
I	+	-	+	-	+	+	+	+
II	+	-	+	-	+	+	-	+
III	+	-	+	-	-	-	-	+
IV	+	-	+	-	-	+	+	+
V	+	-	+	-	+	-	+	+
VI	+	-	+	-	-	-	+	+
VII	-	-	+	-	+	+	+	+
VIII	-	-	+	-	-	+	+	+

Todas as versões foram testadas juntamente com o vídeo base, para se entender o impacto dos dois modelos testados em campanhas. Os resultados de *IPM* e *IPM ratio* em relação ao vídeo base obtidos estão no Quadro 24 abaixo.

Quadro 24 - IPM e IPM ratio obtidos nos testes finais

Versões	A	B	C	D	E	F	G	H	<i>IPM</i>	<i>IPM ratio</i>
Base	-	-	-	-	-	-	-	-	5,5930	1,0000
I	+	-	+	-	+	+	+	+	7,7154	1,3795
II	+	-	+	-	+	+	-	+	7.7306	1.3822
III	+	-	+	-	-	-	-	+	7,3776	1,3191
IV	+	-	+	-	-	+	+	+	8,2656	1,4779
V	+	-	+	-	+	-	+	+	8,6282	1,5427
VI	+	-	+	-	-	-	+	+	8.5181	1.5230
VII	-	-	+	-	+	+	+	+	7,3263	1,3099
VIII	-	-	+	-	-	+	+	+	7,6741	1,3721

6.2.6. Comparação final dos modelos

A comparação final dos modelos foi dividida em duas partes: a primeira trata da diferença de custos entre ambos e a segunda da diferença de resultados.

6.2.6.1. *Comparação de custos*

Para a comparação de custo de produção para versões, devido à confidencialidade do custo da mão-de-obra dos artistas e dos custos de operação das campanhas, não se pode utilizar os dados diretos. Porém, utilizando-se o custo por artista-dia (unidade de medida que leva em consideração a produção média de um artista durante um dia útil), além do custo médio de campanha para um criativo durante o período considerado sua vida útil, calculou-se que o valor de uma melhoria de performance de 0,04% corresponde a aproximadamente o custo de um artista-dia.

O tempo de produção para as versões pedidas no Modelo Atual foi de 30,5 artistas-dia, enquanto o tempo de produção para o Experimento foi de 50 artistas-dia. Apesar do Experimento apresentar um custo de produção, portanto, 63,93% superior ao Modelo Atual, isso corresponde a uma necessidade de diferença de performance final de apenas 0,78% para que o custo seja compensado. Como a ordem de grandeza dos resultados obtidos é consideravelmente maior a 0,78%, os custos de produção são considerados indiferentes na comparação dos dois processos.

Já para os custos de teste, apesar de não se poder explicitar, como anteriormente, a porcentagem de melhoria de performance que eles representam devido à confidencialidade do custo de campanhas ativas, tem-se, também, uma melhoria de performance resultante com ordem de grandeza consideravelmente superior ao custo. Sabe-se que o Experimento apresentou um custo de teste 77,78% superior ao do Modelo Atual, mas, assim como o custo de produção, por ser muito inferior ao custo das campanhas ativas e, conseqüentemente, aos ganhos provenientes de uma melhoria de performance, esse fator é considerado indiferente para as conclusões finais.

A irrelevância dos custos de produção e testes ocorre devido aos altos custos diários de campanhas ativas da empresa. Como esse é um dos principais gargalos para a obtenção de novos clientes é, também, um dos principais investimentos da empresa. Além disso, a produção interna de vídeos, com equipe de artistas contratados em período integral pela empresa possibilita a agilidade necessária para que os custos de mão-de-obra sejam suficientemente baixos e absorvidos pelos processos rotineiros da área.

6.2.6.2. Comparação de resultados

Primeiramente, levando-se em consideração apenas os efeitos individuais dos fatores, obteve-se, para ambos os modelos, a mesma versão final (com os fatores A, C, E, F, G e H). Isso mostra que a forma atual de testes utilizada na empresa, apesar de sua simplicidade, apresenta resultados consistentes. Para o vídeo e as iterações estudadas, o resultado direto do Modelo Atual e da avaliação apenas dos efeitos individuais do Experimento resultaram em uma melhoria de 37,95% de performance.

Para o Modelo Atual, os resultados obtidos para a versão feita diretamente a partir dos dados (versão I) e para as duas versões extra pedidas (versões II e III) estão mostrados na Figura 25 abaixo. Nota-se que a exclusão do fator G (cena de *upgrade* do monstro de lama) teve um impacto positivo, porém pequeno em relação à versão I, o que reforça a hipótese de que duas pré-cartelas em um mesmo vídeo podem prejudicar sua performance. Já a versão feita apenas com os fatores com maiores impactos obtidos teve um resultado consideravelmente inferior ao da versão I.

Resultados finais das versões consequentes da análise do Modelo Atual

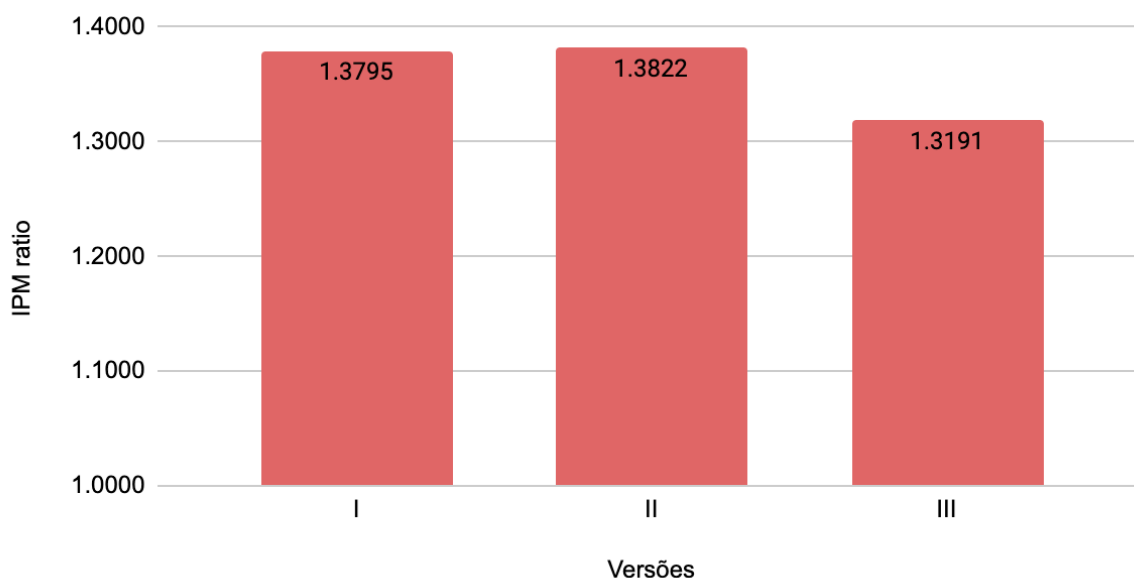


Figura 25 - Resultados finais das versões consequentes da análise do Modelo Atual

A diferença de performance da versão I para a III foi de 6,04%, o que é consideravelmente superior à diferença de custo de produção entre as duas: 1,5

artistas-dia (equivalente a uma diferença de 0,06% em performance). Nesses casos, a vantagem de se produzir com menos alterações ocorre apenas para casos em que a diferença de performance entre as versões é extremamente baixa. Isso seria esperado quando os efeitos das interações entre os fatores são suficientemente negativos para que a presença de muitos deles não seja tão vantajosa. Como a interação entre fatores não é obtida diretamente pelo Modelo Atual, esse tipo de teste é comum na empresa.

Levando-se em consideração, agora, as interações entre efeitos obtidas no Experimento, tem-se que, para o conjunto de interações com maior impacto ($l_{AE} = l_{BC} = l_{DH} = l_{FG}$), foram pedidas as versões IV, V e VI. A Figura 26 mostra os resultados obtidos por elas, além daqueles da versão I.

Resultados finais das versões consequentes da análise do primeiro conjunto de interações do Experimento

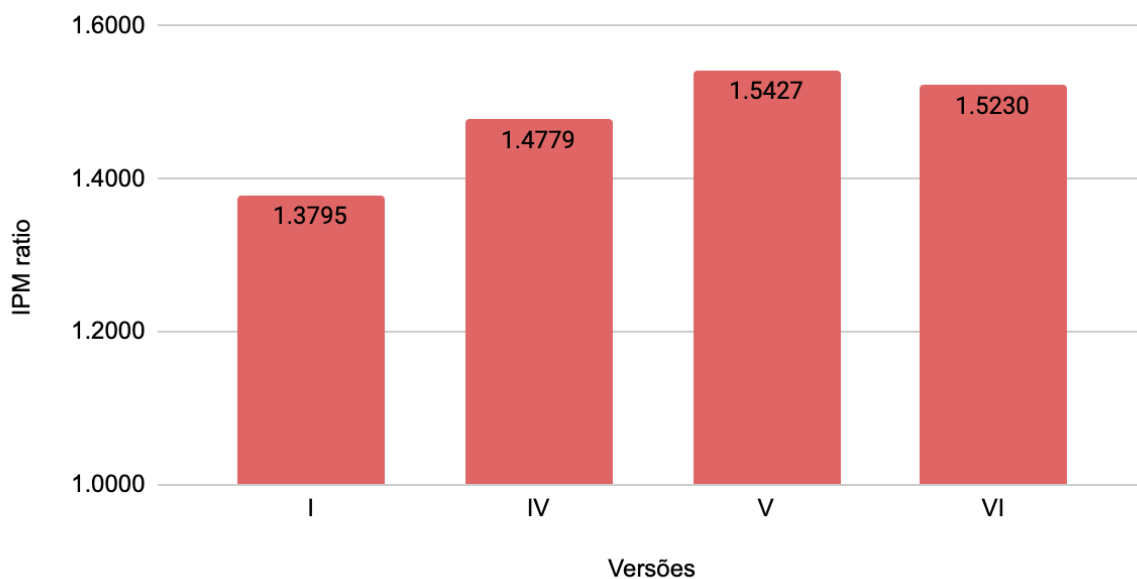


Figura 26 - Resultados finais das versões consequentes da análise do primeiro conjunto de interações do Experimento

Nota-se que ao se excluir o fator E (cena com destaque para a carta), a nova versão (IV) teve um resultado superior à versão I em 9,84%. Isso é inesperado devido ao efeito positivo de 3,16% desse fator. A explicação para esse acontecimento pode estar em algum erro proveniente de variações nos testes ou em alguma interação negativa não considerada envolvendo a alteração E e outra presente na versão I. A

principal hipótese criada sobre isso é a de que o efeito da interação entre os fatores C e E deveria ter sido levado em consideração. Ele apresentou o terceiro maior impacto negativo em performance, de 2,07%. Apesar de não estar acima da margem de erro de 2,40%, esse valor encontra-se bem próximo. Além disso, a interação C (adição de *HUD*) tem como objetivo dar ao espectador uma sensação maior de jogabilidade do vídeo, ou seja, de que o que se está vendo é jogável. Já a adição da cena com destaque para a carta (fator E) possui efeito contrário: ela dá ao vídeo uma característica mais cinematográfica e distancia mais o vídeo de um jogo real. Esse tipo de contradição é, normalmente, evitado na produção e, apesar de não ter obtido um resultado suficientemente relevante, a interação negativa entre os fatores C e E é esperada e é o que se suspeita ter sido a causa dessa diferença de performance entre versões.

A exclusão do fator F (cena de *upgrade* do cavaleiro) na versão V também apresentou resultado superior à versão I em performance. Espera-se que esse resultado seja consequência da interferência negativa que a presença de duas pré-cartelas em um mesmo vídeo pode possuir.

Já na versão VI, em que foram excluídos ambos os fatores E e F, o resultado foi ligeiramente inferior ao da versão V. Pela hipótese formulada para a melhoria de performance da versão IV em relação à I, essa versão deveria ter apresentado uma performance superior à da V. Mais uma vez, essa diferença em relação ao que se esperava pode ter se dado por variações pontuais no teste realizado ou por alguma outra interação não considerada (possivelmente com 3 ou mais fatores envolvidos). Porém, dado que a diferença de performance entre as versões V e VI foi de apenas 1,97%, essa diferença não deve ser considerada relevante e pode-se considerar um empate.

Para as versões produzidas a partir das análises do segundo conjunto de interações ($l_{AG} = l_{BH} = l_{CD} = l_{EF}$), as variações de *IPM* obtidas em relação ao vídeo base, comparadas àquela obtida pela versão I, encontram-se na Figura 27 abaixo.

Resultados finais das versões consequentes da análise do segundo conjunto de interações do Experimento

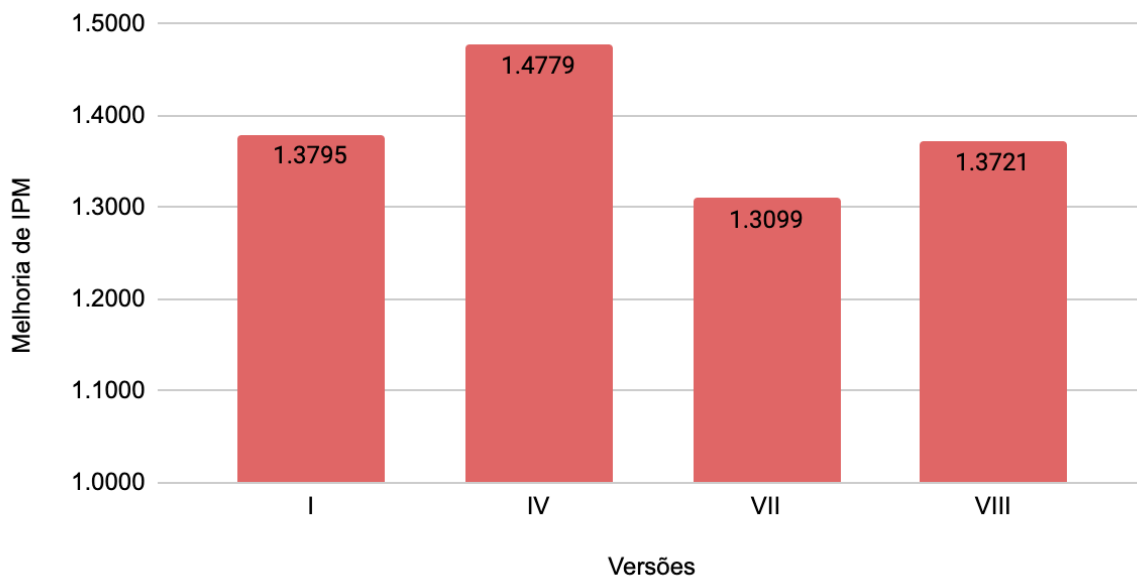


Figura 27 - Resultados finais das versões consequentes da análise do segundo conjunto de interações do Experimento

Nota-se que, para os casos em que se excluiu o fator A (posicionamento dos esqueletos), em ambas as versões VII e VIII, houve uma queda significativa de performance, de 6,96% da versão I para a VII e de 10,58% da versão IV para a VIII. Isso provavelmente se dá devido ao elevado efeito positivo esperado pela presença do fator A ($l_A = 7,73\%$) e reforça a hipótese de que a interação com efeito negativo mais significativo nesse conjunto não é a que envolve a alteração A.

Além disso, assim como visto anteriormente, em ambos os casos em que se excluiu o fator E (cena com destaque para a carta) ocorreu uma melhoria da performance. Como em ambos os casos existe a presença, também, do fator C (HUD), é reforçada a hipótese de que a interação entre as duas alterações possui um efeito negativo relevante na performance.

Comparando-se, agora, o melhor resultado obtido em cada um dos modelos, temos que em um processo comum realizado na empresa, seria atingida uma melhoria de performance de 38,22% (versão II) em relação ao vídeo base, enquanto, ao se fazer o experimento proposto no Experimento, seria obtida uma melhoria de performance de 54,27% (versão V). A diferença de performance obtida, portanto, foi de 16,05% e, como os custos são considerados irrelevantes, conclui-se que o Experimento é melhor para os resultados da empresa e deveria ser adotado.

As principais vantagens observadas pelo Experimento, além da clara visibilidade para os efeitos de interações entre fatores, que forneceram uma fonte de dados concreta e confiável para as versões extra pedidas (que se mostraram importantes para conclusões obtidas), são que os efeitos para cada fator são calculados a partir de duas médias de oito testes cada. Isso fornece ao Experimento uma confiabilidade maior em relação ao Modelo Atual, que chega ao efeito de cada fator a partir da média de apenas dois resultados. Isso pode ter sido o motivo para que no Modelo Atual, a cena de *upgrade* do cavaleiro (fator F) foi considerada superior à do *upgrade* do monstro de lama (fator G) e no Experimento encontrou-se o oposto. Nos resultados obtidos para as versões II e V, em que se exclui, respectivamente, os fatores G e F, tem-se como única variação a alteração da pré-cartela e obteve-se uma diferença de performance de 16,05%. Esse resultado pode ter sido consequência de uma variação entre os testes, acredita-se que o *upgrade* do monstro de lama tenha, realmente, uma performance superior ao *upgrade* do cavaleiro. Uma hipótese que se considera importante para o jogo *Castle Crush* na empresa é a de que vídeos que ressaltam a variedade de personagens no jogo possuem um resultado superior. Dado que na cena de *upgrade* do monstro de lama são mostrados outros personagens, e na do cavaleiro apenas varia-se a arma utilizada, é possível que isso torne o fator G melhor que o fator F para a performance.

7. CONCLUSÕES E PRÓXIMOS PASSOS

Para essa monografia foi proposto um modelo baseado em um experimento fatorial fracionado 2^{k-p} como alternativa ao atual processo realizado para conclusões a respeito de hipóteses em relação a iterações para anúncios. Baseando-se nos custos e resultados obtidos para a realização dos processos completos dos dois modelos estudados: o atualmente aplicado e o proposto, concluiu-se que a utilização de experimentos fatoriais fracionados 2^{k-p} para a testagem de hipóteses sobre iterações de vídeos é preferível para a Wildlife Studios.

Dentre as desvantagens do modelo sugerido, tem-se que ele exige um preparo mais elaborado, em que oito hipóteses diferentes devem estar estabelecidas previamente e devem ser combinadas corretamente para que se alcance os resultados esperados. Além disso, os custos de produção e de testagem desse modelo são consideravelmente superiores.

Essas desvantagens, porém, são consideradas pouco relevantes devido ao alto custo diário das campanhas ativas e, portanto, alto valor conquistado por otimizações de criativos. A visibilidade dada aos efeitos de interações de fatores possibilitada pelo delineamento de experimento proposto fornece um fundamento matemático para versões extras pedidas, que contrasta com os conhecimentos da área passados informalmente entre as pessoas e sem uma comprovação teórica obtida. Além disso, ao calcular-se o efeito das iterações baseado em duas médias de oito testes cada, em contraste com uma média entre dois testes realizada no modelo atual, obtém-se uma confiança muito maior em relação aos resultados e pode-se obter conclusões mais concretas a respeito da performance dos fatores avaliados.

Esse estudo será apresentado aos responsáveis pela testagem de hipóteses na área de *Creatives Intelligence* e, provavelmente, será replicado mais vezes para se garantir sua funcionalidade antes da sua aplicação efetiva nos processos da empresa.

8. REFERÊNCIAS

TANG, A. K. Y. Electronic Commerce Research and Applications. **Elsevier**, Hong Kong, 2019.

TAKAHASHI, D. SuperData: Games hit \$120.1 billion in 2019, with Fortnite topping \$1.8 billion. **VentureBeat**, January 2020.

RUTZ, O.; ARAVINDAKSHAN, A.; RUBEL, O. Measuring and forecasting mobile game app engagement. **Elsevier**, Janeiro 2019.

GAGLIONI, C. A ascensão do universo dos games. E sua potência no século 21. **Nexo Jornal**, Setembro 2019.

AGÊNCIA BRASIL. Como estão sobrevivendo os games mobile em meio à quarentena? **Istoé**, n. Dinheiro, Abril 2020.

PRNEWswire. Fundada por brasileiros, Wildlife Studios é avaliada em US\$ 1.3 bi e alcança patamar de unicórnio. **Exame**, 06 Dezembro 2019.

GRANATO, L. Unicórnio de games abre mais de 100 vagas no Brasil e na Argentina. **Exame**, 03 Janeiro 2020.

WILDLIFE STUDIOS. About Us. **Wildlife Studios**, 2020. Disponível em: <<https://wildlifestudios.com/about-us/>>. Acesso em: 3 Maio 2020.

BOX, G. E. P.; HUNTER, J. S.; HUNTER, W. G. **Statistics for Experimenters - Design, Innovation and Discovery**. Second Edition. ed. Hoboken, New Jersey: John Wiley & Sons, 2005.

FERREIRA, P. V. **Introdução à Estatística Experimental**. Campus de Engenharias e Ciências Agrárias da Universidade Federal de Alagoas. Maceió. 2011.

CALEGARE, Á. J. D. A. **Introdução ao Delineamento de Experimentos**. Segunda Edição. ed. São Paulo: Blucher, 2009.

SELTMAN, H. **Approximations for Mean and Variance of a Ratio**. Carnegie Mellon University. Pittsburgh. 2018.

ANTONY, J. **Design of Experiments for Engineers and Scientists**. Segunda Edição. ed. Edimburgo: Elsevier Ltd., 2014.

MONTGOMERY, D. C.; RUNGER, G. C. **Applied Statistics and Probability for Engineers**. Terceira Edição. ed. [S.l.]: John Wiley & Sons Inc., 2003.