



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO ACADÊMICO DO AGRESTE
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE PRODUÇÃO

FLÁVIO LOPES DA SILVA FILHO

**ESTUDO COMPORTAMENTAL DO PROCESSO DE ELICITAÇÃO
INTRACRITÉRIO COM SAD DO FITRADEOFF: uma abordagem utilizando o eye-
tracking**

Caruaru
2024

FLÁVIO LOPES DA SILVA FILHO

**ESTUDO COMPORTAMENTAL DO PROCESSO DE ELICITAÇÃO
INTRACRITÉRIO COM SAD DO FITRADEOFF: uma abordagem utilizando o eye-
tracking**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal de Pernambuco do Centro Acadêmico do Agreste, como requisito parcial para obtenção do título de mestre(a) em Engenharia de produção. Área de concentração: Otimização e Gestão da Produção.

Orientador (a): Anderson Lucas Carneiro de Lima da Silva

Caruaru

2024

.Catalogação de Publicação na Fonte. UFPE - Biblioteca Central

Silva Filho, Flavio Lopes da.

Estudo comportamental do processo de elicitação intracritério com o SAD do FITradeoff: uma abordagem utilizando o eye-tracking / Flavio Lopes da Silva Filho. - Caruaru, 2024.

125f.: il.

Dissertação (Mestrado) - Universidade Federal de Pernambuco, Centro Acadêmico do Agreste, Programa de Pós-Graduação em Engenharia de Produção, 2024.

Orientação: Anderson Lucas Carneiro de Lima da Silva.

Inclui referências.

1. FITradeoff; 2. Decisão multicritério; 3. Eye-tracking; 4. Estudo comportamental. I. Silva, Anderson Lucas Carneiro de Lima da. II. Título.

UFPE-Biblioteca Central

FLÁVIO LOPES DA SILVA FILHO

**ESTUDO COMPORTAMENTAL DO PROCESSO DE ELICITAÇÃO
INTRACRITÉRIO COM SAD DO FITRADEOFF: uma abordagem utilizando o eye-
tracking**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal de Pernambuco do Centro Acadêmico do Agreste, como requisito parcial para obtenção do título de mestre(a) em Engenharia de produção. Área de concentração: Otimização e Gestão da Produção.

Aprovado em: 26 / 02 / 2025.

BANCA EXAMINADORA

Prof. Dr. Anderson Lucas Carneiro de Lima da Silva (Orientadora)
Universidade Federal de Pernambuco - UFPE

Prof^a. Dr^a. Thárcylla Rebecca Negreiros Clemente (Examinador Interno)
Universidade Federal de Pernambuco - UFPE

Prof^a. Dr^a. Mariana Rodrigues de Almeida (Examinador Externo)
Universidade Federal do Rio Grande do Norte - UFRN

Dedico este trabalho a todos que fizeram, fazem e farão parte desta caminhada. Aos que me apoiaram nos momentos difíceis, comemoraram comigo cada conquista e, de alguma forma, contribuíram para que este momento se tornasse possível. Esta vitória é nossa.

AGRADECIMENTOS

Primeiramente, agradeço a Deus, por me conceder sabedoria e discernimento ao longo dessa caminhada, permitindo-me superar os desafios e alcançar este importante momento. Aos meus pais, Maria Zélia Florêncio Lopes e Flávio Lopes da Silva, minha eterna gratidão pelo apoio incondicional e pelo incentivo constante a seguir no caminho da educação. Sem o amor e a dedicação de vocês, nada disso teria sido possível.

A minha tia, Maria Tereza Alves Florêncio, por sempre acreditar em mim e me oferecer seu carinho. A minha namorada, Carolina Azevedo, por sua paciência, compreensão e apoio constante durante toda essa jornada. Sua presença foi fundamental para o meu equilíbrio e motivação.

Aos meus colegas de mestrado, que tornaram essa caminhada mais leve e cheia de aprendizados. Obrigado por todo apoio, trocas de ideias e pela amizade.

Ao meu orientador, Anderson Lucas Carneiro Lima da Silva, minha imensa gratidão pela paciência, orientação e pelos valiosos ensinamentos que me proporcionou ao longo deste ciclo. Seu conhecimento e dedicação foram essenciais para o desenvolvimento deste trabalho.

Aos meus colegas de trabalho, que me deram suporte nos momentos de ausência, em especial à Joana Darc Cintra, por toda confiança e apoio que me deu ao longo dessa trajetória. Sua colaboração foi crucial para a conclusão deste projeto.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

Por fim, agradeço a todos que, de forma direta ou indireta, contribuíram para que este objetivo fosse alcançado. Cada gesto, palavra ou apoio foi fundamental para que este momento se tornasse realidade.

RESUMO

Este trabalho tem como objetivo compreender o comportamento do decisor, na avaliação intracritério do método FITradeoff, utilizando ferramentas da neurociência como meio de obter informações sobre o esforço cognitivo empregado e a atenção dos usuários, durante o processo decisório. O estudo abordou o uso do eye-tracking como forma de captar os aspectos cognitivos subjacentes do decisor, ao longo da etapa de avaliação intracritério do SAD do FITradeoff. Esta obra analisou quatro grupos, definidos pelo gênero e nível de conhecimento prévio dos participantes do experimento, buscando identificar padrões comportamentais. Foi analisado o tempo que cada participante levou para concluir as rodadas de avaliação intracritério e a influência que cada rodada tem sobre a outra. Outro ponto abordado nessa dissertação, foi a relação que o gênero e o conhecimento prévio em métodos multicritério dos participantes têm sobre o tempo de foco nas áreas de interesse previamente definidas, assim, investigando se participantes com perfis semelhantes focam sua atenção nas mesmas áreas de interesse e oferecendo, dessa maneira, suporte aos analistas através dos padrões encontrados em cada perfil. Essa obra ainda se propõe a contribuir com os desenvolvedores do SAD do FITradeoff, a fim de melhorar a experiência do usuário, reduzindo os elementos da interface que podem causar distrações e possam interferir na qualidade dos resultados, dessa maneira, ajustando-se a diferentes perfis e promovendo o aumento na eficácia do software.

Palavras-chave: FITradeoff, Decisão multicritério, Eye-tracking, Estudo Comportamental.

ABSTRACT

This work aims to understand the behavior of the decision-maker in the intra-criteria evaluation of the FITradeoff method, using neuroscience tools to gather information about the cognitive effort employed and the attention of users during the decision-making process. The study addressed the use of eye-tracking as a way to capture the cognitive aspects underlying the decision-maker, throughout the intra-criteria evaluation step of the FITradeoff SAD. This work analyzed four groups, defined by the gender and prior knowledge level of the experiment participants, aiming to identify behavioral patterns. It analyzed the time each participant took to complete the intra-criteria evaluation rounds and the influence that each round has on the others. Another point discussed in this dissertation was the relationship between the participants' gender and prior knowledge in multi-criteria methods and the focus time on predefined areas of interest, investigating whether participants with similar profiles focus their attention on the same areas of interest, thus offering support to analysts through the patterns found in each profile. This work also aims to contribute to the developers of the FITradeoff SAD, in order to improve the user experience by reducing elements in the interface that may cause distractions and interfere with the quality of results, thus adjusting to different profiles and enhancing the software's effectiveness.

Keywords: FITradeoff; Multi-criteria decision; Eye tracking; Behavioral Study.

LISTA DE ILUSTRAÇÕES

Figura 1- Estrutura desse trabalho	19
Figura 2 - Etapas do experimento	32
Figura 3 - Aparelhos de eletroencefalograma utilizado no experimento	33
Figura 4 - Cruz de fixação para calibração do Eye-tracking	34
Figura 5 - Calibração por perseguição de pontos aleatórios	34
Figura 6 - Tela de Login da página do FITradeoff	35
Figura 7 - Escolha da problemática	35
Figura 8 - Interface do FITradeoff na etapa intreacritério	38
Figura 9 - Seleção de critérios na avaliação intracritério	38
Figura 10 - Avaliação dos critérios na interface FITradeoff	39
Figura 11 - Declaração de valor linear no SAD FITradeoff.....	40
Figura 12 - Áreas de interesse	41
Figura 13 - Diagrama Boxplot para a amostra	47
Figura 14 - Diagrama Boxplot para o grupo homens com conhecimento	49
Figura 15 - Tempo média de duração das rodadas para o grupo homens com conhecimento .	51
Figura 16 - Diagrama boxplot para o grupo mulheres com conhecimento.....	52
Figura 17 - Tempo média de duração das rodadas para o grupo mulheres com conhecimento	53
Figura 18 - Diagrama Boxplot para o grupo homens sem conhecimento.....	55
Figura 19 - Tempo média de duração das rodadas para o grupo homens sem conhecimento .	56
Figura 20 - Diagrama Boxplot para o grupo mulheres sem conhecimento.....	57
Figura 21 - Tempo média de duração das rodadas para o grupo mulheres sem conhecimento	59
Figura 22 - Comparação entre as médias das rodadas dos grupos	60
Figura 23 - Teste de esfericidade para o grupo homens com conhecimento.	65
Figura 24 - Resultado da ANOVA corrigido para grupo homens com conhecimento.....	66
Figura 25 - Teste de esfericidade para o grupo mulheres com conhecimento.	69
Figura 26 - Resultado da ANOVA corrigido para grupo mulheres com conhecimento.....	69
Figura 27 - Teste de esfericidade para o grupo homens sem conhecimento.....	72
Figura 28 - Resultado da ANOVA para grupo homens sem conhecimento.	72
Figura 29 - Teste de esfericidade para o grupo mulheres sem conhecimento.....	75
Figura 30 - Resultado da ANOVA para grupo mulheres sem conhecimento.	76
Figura 31 - Teste estatístico para comparar as rodadas.	78
Figura 32 - Visualização gráfica da matriz de correlação.....	81
Figura 33 - Curva de aprendizagem a amostra.....	83
Figura 34 - Curva de aprendizagem da amostra completa.....	84
Figura 35 - Curva de aprendizagem ajustada grupo homens com conhecimento.	85
Figura 36 - Curva ajustada dos participantes do grupo homens com conhecimento.....	86
Figura 37 - Curva de aprendizagem ajustada para o grupo mulheres com conhecimento.....	87
Figura 38 - Curva ajustada dos participantes do grupo mulheres com conhecimento.	88
Figura 39 - Curva de aprendizagem ajustada grupo homens sem conhecimento.....	89
Figura 40 - Curva ajustada dos participantes do grupo homens sem conhecimento.....	90
Figura 41 - Curva ajustada dos participantes do grupo mulheres sem conhecimento.....	91
Figura 42 - Curva ajustada dos participantes do grupo mulheres sem conhecimento.....	92
Figura 43 - Histograma da taxa de aprendizagem.	94

Figura 44 - Gráfico dos resíduos para a taxa de aprendizagem.....	94
Figura 45 - Matriz de interação entre gênero e conhecimento.....	99
Figura 46 - Impacto da permutação das variáveis.....	100
Figura 47 - Importância das variáveis decompostas.....	101
Figura 48 - Comparação da fixação na área Choice.....	105
Figura 49 - Comparação da fixação na área Graph.....	106
Figura 50 - Comparação da fixação na área InstB.....	107
Figura 51 - Comparação da fixação na área InstD.....	108
Figura 52 - ANOVA da área Choice na primeira rodada.....	109
Figura 53 - ANOVA da área Graph na primeira rodada.....	110
Figura 54 - ANOVA da área Choice na segunda rodada.....	110
Figura 55 - ANOVA da área Choice na terceira rodada.....	111
Figura 56 – ANOVA da área Choice na quarta rodada.....	111
Figura 57 - ANOVA da área Choice na setima rodada.....	112
Figura 58 - Resumo dos resultados obtidos nesse trabalho.....	113

LISTA DE TABELAS

Tabela 1 - Alternativas e critérios.....	36
Tabela 2 – Grupos que compõem a amostra	37
Tabela 3 - Matriz de decisão	37
Tabela 4 - Descrição das áreas de interesse	41
Tabela 5 – Duração, em segundos, da elicitação de cada critérios, para todos os participantes	45
Tabela 6 - Estatísticas descritivas da Amostra.	46
Tabela 7 - Estatísticas descritivas do grupo homens com conhecimento	49
Tabela 8 - Estatísticas descritivas do grupo mulheres com conhecimento.	51
Tabela 9 - Estatísticas descritivas do grupo homens sem conhecimento.....	54
Tabela 10 - Tipos de ANOVA e seus correspondentes não paramétricos.	62
Tabela 11 - Teste de normalidade e homocedásticos para o grupo Homem com conhecimento	63
Tabela 12 - Teste de normalidade e Homoscedasticidade para o grupo mulheres com conhecimento.....	63
Tabela 13 - Teste de normalidade e homoscedasticidade para o grupo Homem sem conhecimento.....	64
Tabela 14 - Teste de normalidade e homoscedasticidade para o grupo mulher sem conhecimento.....	64
Tabela 15 - Teste de contraste para o grupo homens com conhecimento.....	66
Tabela 16 - Teste pós-hoc para o grupo homens com conhecimento.	67
Tabela 17 - Teste de contraste para o grupo mulheres com conhecimento.....	69
Tabela 18 - Teste pós-hoc para o grupo mulheres com conhecimento.....	70
Tabela 19 - Teste de contraste para o grupo homens sem conhecimento	73
Tabela 20 - Teste pós-hoc para o grupo homens sem conhecimento.	74
Tabela 21 - Teste de contraste para o grupo mulheres sem conhecimento	76
Tabela 22 - Teste de normalidade para amostra.	79
Tabela 23 - Matriz de correlação para amostra.	79
Tabela 24 - Descrição dos valores de correlação.....	80
Tabela 25 - Melhor e pior relação entre as rodadas.	80
Tabela 26 - Valor da estatística t.	82
Tabela 27 - p-valor associado a estatística t.....	82
Tabela 28 - Testes de normalidade e homoscedasticidade para a taxa de aprendizado.....	93
Tabela 29 - Teste t para os grupos homens com conhecimento e homens sem conhecimento.....	95
Tabela 30 - Teste t para os grupos mulheres com conhecimento e mulheres sem conhecimento.....	95
Tabela 31 - Teste t para os grupos homens com conhecimento e mulheres com conhecimento	95
Tabela 32 - Teste t para os grupos homens com conhecimento e mulheres sem conhecimento.	96
Tabela 33 - Teste t para os grupos homens com conhecimento e mulheres sem conhecimento.	96
Tabela 34 - Teste para os grupos homens sem conhecimento e mulheres sem conhecimento.....	97

Tabela 35 – Teste t para os grupos homens com conhecimento e Mulheres sem conhecimento.	98
Tabela 36 - Importância das variáveis para o modelo	100
Tabela 37 - Porcentagem do impacto de cada variável decomposta	101
Tabela 38 - Tempo médio de fixação do grupo homens com conhecimento.....	102
Tabela 39 - Tempo médio de fixação do grupo mulheres com conhecimento.....	103
Tabela 40 - Tempo médio de fixação do grupo homens sem conhecimento.	104
Tabela 41 - Tempo médio de fixação do grupo mulheres sem conhecimento.	104

LISTA DE QUADROS

Quadro 1 - Erros do Tipo 1 e Tipo 11	61
Quadro 2 - Pressupostos da ANOVA	61
Quadro 3 - Teste de Turkey para comparação entre os grupos.	77
Quadro 4 - Decomposição das variáveis.	100

LISTA DE ABREVIATURAS E SIGLAS

AOI	Área de Interesse
EEG	Eletroencefalograma
ANOVA	Análise de variância
SAD	Sistemas de Apoio a Decisão

SUMÁRIO

1.	Introdução	14
1.1.	Justificativa	15
1.2.	Objetivos	17
1.2.1.	Objetivo Geral.....	17
1.2.2.	Objetivos Específicos	17
1.3.	Metodologia	17
1.4.	Estrutura do trabalho.....	20
2.	Referencial Teórico e revisão da literatura	21
2.1.	Referencial Teórico	21
2.1.1.	FITradeoff	21
2.1.2.	Neurociência para Decisão.....	24
2.2.	Revisão da Literatura	26
2.2.1.	Estudos de Neurociência para Decisão.....	26
2.2.2.	Estudos de Neurociência com o FITradeoff	28
2.2.3.	Síntese do Estado da Arte e Posicionamento deste Trabalho	31
3.	Experimento	32
3.1.	Problema de decisão.....	36
3.2.	Amostras	36
3.3.	Avaliação intracritério.....	37
3.4.	Eye-Tracking	40
3.5.	Estatísticas.....	42
4.	Resultados e Discussões.....	44
4.1.	Curva de Aprendizagem	44
4.1.1.	Estatísticas descritiva	46
4.1.1.1.	<i>Estatística descritiva do grupo homens com conhecimento</i>	<i>48</i>
4.1.1.2.	<i>Estatística descritiva do grupo mulheres com conhecimento</i>	<i>51</i>
4.1.1.3.	<i>Estatística descritiva do grupo homens sem conhecimento</i>	<i>54</i>
4.1.1.4.	<i>Estatística descritiva do grupo mulheres sem conhecimento</i>	<i>56</i>
4.1.1.5.	<i>Comparação entre os grupos</i>	<i>59</i>
4.1.2.	Testes paramétricos e não paramétricos.....	60
4.1.2.1.	<i>Grupo homens com conhecimento</i>	<i>65</i>

4.1.2.2.	<i>Grupo mulheres com conhecimento</i>	68
4.1.2.3.	<i>Grupo homens sem conhecimento</i>	72
4.1.2.4.	<i>Grupo mulheres sem conhecimento</i>	75
4.1.2.5.	<i>Comparação entre os grupos</i>	76
4.1.3.	Análise de correlação	78
4.1.4.	Curva de aprendizagem	83
4.1.4.1.	<i>Grupo homens com conhecimento</i>	85
4.1.4.2.	<i>Grupo mulheres com conhecimento</i>	87
4.1.4.3.	<i>Grupo homens sem conhecimento</i>	88
4.1.4.4.	<i>Grupo mulheres sem conhecimento</i>	90
4.1.4.5.	<i>Comparação entre grupos</i>	92
4.1.5.	Abordagem Random Forest	98
4.2.	Tempo de fixação	102
4.2.1.	Análise descritiva	102
4.2.2.	Análise de variância para o tempo de fixação	108
4.2.2.1.	Rodada 1	109
4.2.2.2.	Rodada 2	110
4.2.2.3.	Rodada 3	110
4.2.2.4.	Rodada 4	111
4.2.2.5.	Rodada 5	112
4.2.2.6.	Rodada 6	112
4.2.2.7.	Rodada 7	112
5.	Conclusão e sugestões para trabalhos futuros	113
5.1.	Conclusões	113
5.2.	Insights e Sugestões de Melhoria	115
5.3.	Limitações e Sugestões de Trabalhos Futuros	116
	REFERÊNCIAS	118

1. INTRODUÇÃO

Com o passar do tempo, o ambiente está cada vez mais cercado de tecnologias e sistemas digitais, essa revolução tecnológica aconteceu rápido e em pouco tempo, e ficou conhecida como a era da informação (CASTELLS, 1996). Nunca se teve acesso a um volume tão grande de informação, vindos de diferentes fontes, em tempo real. As informações que antes demoravam dias ou semanas para serem processadas agora estão disponíveis em questão de segundos (BRYNJOLFSSON *et al*, 2014). Esse cenário transformou a informação no recurso mais precioso do mundo globalizado. Ela não é apenas uma ferramenta, mas sim a base para inovação, tomada de decisões e sobrevivência no mercado. As organizações que conseguem lidar bem com o crescente volume de dados e transformá-los em conhecimento têm uma vantagem competitiva e estratégica sobre seus concorrentes. No entanto, o simples acúmulo de dados não é suficiente. O verdadeiro diferencial está na capacidade de compreender esses dados, extrair insights e utilizá-los de forma precisa e eficiente (MAYER-SCHÖNBERGER *et al*, 2013).

Por outro lado, essa abundância de informações também trouxe desafios significativos, especialmente no que diz respeito à tomada de decisões. No mundo globalizado em que vivemos, as decisões não podem mais ser baseadas apenas em intuição ou experiência acumulada. Elas precisam ser rápidas, embasadas em análises confiáveis e, ao mesmo tempo, levar em conta a complexidade dos contextos em que estão inseridas. Além disso, as decisões de hoje frequentemente envolvem múltiplos critérios e objetivos conflitantes, o que as torna ainda mais desafiadoras. Assim, surgem as ferramentas de apoio à decisão como aliadas indispensáveis para lidar com essas complexidades, permitindo que informações sejam organizadas e analisadas de maneira eficiente, auxiliando na escolha de caminhos mais racionais e estratégicos (KEENEY & RAIFFA, 1993).

Nos últimos anos, a neurociência tem demonstrado um enorme potencial para melhorar esses sistemas ao permitir a captação de aspectos cognitivos subjacentes das preferências humanas. Ferramentas como rastreamento ocular (*eye tracking*), eletroencefalografia (EEG) e análise de respostas fisiológicas estão sendo cada vez mais utilizadas para explorar como as pessoas processam informações e tomam decisões (ARIELY & BERNS, 2010). Essas ferramentas vão além do que as pessoas conseguem verbalizar, revelando aspectos como a influência das emoções, do gênero, do conhecimento prévio e dos vieses cognitivos, que muitas vezes passam despercebidos e comumente são generalizados e tratados de forma semelhante.

O método de Tradeoff Flexível e Interativo (FITradeoff) (De Almeida *et al.*, 2016; De Almeida *et al.*, 2021; De Almeida *et al.*, 2014) já é reconhecido por sua eficiência em facilitar decisões multicritério, com uma abordagem que reduz a carga cognitiva do decisor e oferece resultados consistentes (DE ALMEIDA *et al.*, 2016). Ainda assim, a integração de dados neurocientíficos pode ampliar significativamente o alcance e a precisão desse modelo. Ao analisar, por exemplo, os padrões de fixação ocular em áreas de interesse e como os grupos evoluem em termos de aprendizado e compreensão ao longo do processo, é possível ajustar o sistema para refletir melhor as preferências e comportamentos reais dos decisores, bem como a condução do processo entre o decisor e o analista de decisão, ator fundamental para suporte metodológico.

Essa dissertação propõe uma análise comportamental, utilizando ferramentas da neurociência, sobre o processo de avaliação intracritério com sistema de apoio a decisão (SAD) baseado no método FITradeoff a partir de diferentes grupos de decisores. A combinação de ferramentas como a análise de fixações oculares e curvas de aprendizagem, permitem não apenas compreender melhor o comportamento dos grupos, mas também aprimorar a experiência do usuário. Como resultado se tem um processo mais personalizado, mais eficiente e com menor risco de inconsistências. Ao longo do trabalho, será discutido como essas interseções entre neurociência e SADs podem ajudar a entender e aperfeiçoar o entendimento sobre como o aprendizado no processo de apoio a decisão multicritério acontece, e como é influenciado por aspectos como o gênero e grau de conhecimento dos participantes.

1.1. Justificativa

A justificativa para a realização deste trabalho está na crescente necessidade de aprimorar os SADs, em um mundo cada vez mais complexo e dinâmico. Para Turban *et al.* (2018), o aumento da quantidade de informações disponíveis e a diversidade de opções que os decisores enfrentam, torna-se essencial que os sistemas utilizados para auxiliar na tomada de decisões sejam não apenas eficazes, mas também intuitivos e adaptáveis às necessidades dos usuários e especialistas. A eficiência de um SAD não se limita apenas à apresentação de dados, mas envolve a capacidade de entender e prever o comportamento dos usuários, facilitando a escolha de alternativas que melhor atendam às suas preferências e objetivos.

Nesse contexto, a neurociência surge como uma ferramenta para o aprimoramento dos SAD. Ao estudar os processos cognitivos e comportamentais que influenciam a tomada de decisão, a neurociência fornece insights sobre como os indivíduos processam informações,

avaliam opções e formam preferências. Para Cohen e Ranganath (2007), a integração de métodos neurocientíficos, como o rastreamento ocular e a eletroencefalografia, permite capturar dados sobre a atenção, o foco e as reações durante a interação com os sistemas. Esses dados ajudam a entender o comportamento dos usuários, possibilitando ajustes na interface e na apresentação das informações de forma mais assertiva e eficiente.

A análise do aprendizado ao longo da utilização dos SADs é fundamental para entender como os usuários se adaptam e se tornam mais proficientes na interação com esses sistemas. À medida que os usuários se familiarizam com o funcionamento e a interface do software, é esperado que sua capacidade de tomar decisões melhore. Essa evolução no aprendizado pode ser medida por meio da curva de aprendizagem utilizando o tempo de execução das etapas do sistema. Descrever como o aprendizado se comporta ao longo do tempo permite que os desenvolvedores identifiquem áreas onde os usuários podem enfrentar dificuldades, possibilitando ajustes que melhorem a experiência e a eficácia do sistema. Além disso, essa análise pode revelar padrões de aprendizado que variam entre diferentes grupos de usuários, fornecendo insights valiosos para a personalização do sistema.

Outro aspecto importante é a influência que o gênero e o conhecimento prévio dos usuários nas áreas do SAD onde mantêm a atenção durante a interação com ele. Zhang e Li (2005) mostram que as diferenças entre o gênero e a diferença cognitiva entre homens e mulheres, podem apresentar distinções significativas em seus padrões de atenção e em como processam informações, o que pode impactar diretamente a eficácia dos SADs. Além disso, o nível de conhecimento prévio dos usuários sobre decisão multicritério desempenha um papel importante na forma como eles interagem com o SAD. Usuários mais experientes podem ter um foco de atenção diferente e uma abordagem distinta na avaliação das opções, em comparação com aqueles que possuem menos familiaridade. Analisar essas variáveis permite que os desenvolvedores criem sistemas que não apenas atendam a um público, mas que também sejam sensíveis às necessidades de grupos específicos, resultando em uma experiência de decisão mais eficaz. Essa compreensão do aprendizado e das influências individuais são essenciais para o desenvolvimento dos SAD, em especial o FITradeoff, de forma que potencializem a capacidade do software de captar as preferências do usuário e convertê-las na melhor alternativa possível, dentro do conjunto apresentado.

Dessa maneira, o trabalho se justifica através da necessidade de desenvolver SAD que não apenas apresentem informações, mas que também compreendam e se adaptem ao comportamento dos usuários. A neurociência, ao fornecer uma compreensão mais profunda dos processos de decisão, se torna uma ferramenta indispensável, colaborando na criação de

sistemas mais eficazes, intuitivos e personalizados. A dissertação busca, contribuir para o avanço do FITradeoff, integrando conhecimentos de neurociência e práticas de design centrado no usuário, com o objetivo de melhorar a qualidade dos resultados, nas diversas problemáticas que o método abrange.

1.2. Objetivos

1.2.1. Objetivo Geral

Esse trabalho tem como objetivo principal avaliar os processos cognitivos e comportamentais, sobre o grau de conhecimento em métodos multicritério e o gênero, que são subjacentes ao processo de eliciação de preferências durante a etapa de avaliação intracritério no SAD do FITradeoff.

1.2.2. Objetivos Específicos

- Avaliar a taxa de aprendizado dos participantes ao longo do experimento, em função do grau de conhecimento e do gênero, na avaliação intracritério do SAD do FITradeoff;
- Explicar a influência entre o gênero e o conhecimento na taxa de aprendizagem dos participantes;
- Oferecer recomendações de melhoria para a etapa de avaliação intracritério do SAD FITradeoff;
- Oferecer suporte aos especialistas para a etapa de avaliação intracritério;

1.3. Metodologia

A metodologia deste trabalho foi estruturada em quatro etapas principais: revisão de literatura, aplicação dos experimentos, análise dos resultados e elaboração de conclusões, insights e melhorias.

A primeira etapa consistiu em uma revisão de literatura relacionada ao SAD do FITradeoff e à aplicação de técnicas de neurociência na tomada de decisão. Essa revisão teve como objetivo identificar os principais conceitos, teorias e estudos anteriores que fundamentam a pesquisa. Foram analisados artigos acadêmicos, teses e publicações relevantes que abordam

a interação entre o FITradeoff e metodologias de rastreamento ocular (eye-tracking) e eletroencefalograma (EEG). A revisão permitiu estabelecer um referencial teórico sólido e identificar lacunas na literatura que justificam a realização deste estudo.

Após a revisão da literatura, foi realizada a aplicação do experimento, que envolveu a seleção da amostra, o desenho do experimento e a coleta de dados. A amostra foi composta por estudantes de graduação e pós-graduação em Engenharia de Produção da Universidade Federal de Pernambuco, além de profissionais do mercado, buscando garantir diversidade em termos de conhecimento e experiência. O experimento foi estruturado em nove etapas, incluindo a introdução ao FITradeoff, explanação do experimento e a etapa de avaliação intracritério, onde os participantes foram expostos a diferentes alternativas de decisão. Durante essa fase, foram utilizados equipamentos de eye-tracking para registrar o número de fixações e a duração das fixações dos participantes nas áreas de interesse. Além disso, questionários foram aplicados antes e após a avaliação para coletar dados demográficos e informações sobre o conhecimento prévio dos participantes.

Os dados coletados foram analisados utilizando técnicas estatísticas descritivas e inferenciais. As métricas oculares foram analisadas e transformadas na taxa de aprendizado dos participantes, permitindo identificar padrões de comportamento. A análise inclui a comparação das médias de fixações e duração das fixações entre diferentes grupos, como homens e mulheres, com e sem conhecimento. Também foi realizada uma análise de regressão para investigar a influência do gênero e do conhecimento nas métricas de atenção nas áreas de interesse previamente definidas. Os resultados foram apresentados em tabelas e gráficos, facilitando a visualização das relações entre as variáveis estudadas.

Com base nos resultados obtidos, foram elaboradas conclusões que destacam as principais contribuições do estudo. Os insights gerados a partir da análise dos dados permitiram identificar a relação entre o comportamento visual dos participantes e sua taxa de aprendizado, evidenciando como a atenção se distribui nas áreas de interesse durante a avaliação. Além disso, foram observadas diferenças significativas entre os grupos em termos de fixações e duração, sugerindo que fatores como gênero e conhecimento influenciam a tomada de decisão. A partir dessas conclusões, foram propostas recomendações de melhorias para o SAD, com o intuito de tornar a interface mais intuitiva e personalizada, aumentando a eficácia do FITradeoff na elicitación de preferências.

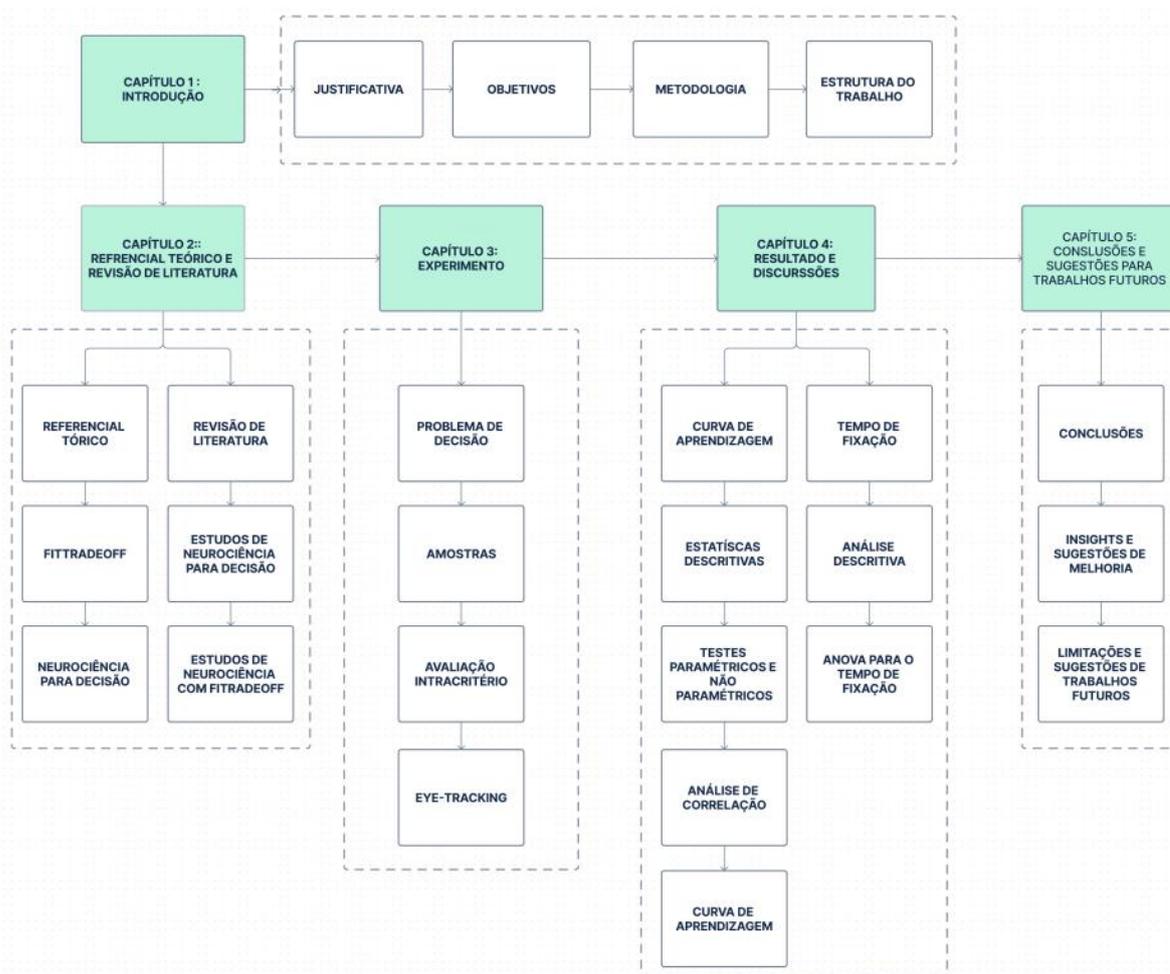
Dessa forma, a metodologia apresentada fornece uma estrutura clara e lógica para a condução da pesquisa, desde a revisão da literatura até a aplicação dos experimentos e a análise dos resultados. Os *insights* e melhorias sugeridos contribuirão para o avanço do conhecimento

na área de apoio à decisão e na aplicação de técnicas de neurociência, oferecendo um caminho para futuras investigações.

Para realizar essa pesquisa, foi utilizado o Tobii Studio para tratar os dados referentes ao rastreamento ocular, além da utilização de softwares para análises estatísticas. Outra ferramenta importante foi a linguagem Python e suas bibliotecas de análises de dados, com a Numpy e Pandas, além das bibliotecas de gerações gráficas como a matplotlib e a da biblioteca Scikit-learn para modelos de random forest.

A Figura 1 apresenta a estrutura desse trabalho, mostra as sessões e subseções, assim como os assuntos tratados.

Figura 1- Estrutura desse trabalho



Fonte: O autor (2024).

1.4. Estrutura do trabalho

Esse trabalho se organiza em 5 capítulos, com suas respectivas sessões e subseções. Cada capítulo traz os conceitos necessários para o entendimento, discussão e interpretação dos resultados.

O primeiro capítulo, traz a parte introdutória desse trabalho, ambientando os leitores sobre as ideias, o tema e o contexto. Ainda nesse capítulo são apresentadas a justificativa, que apresenta o motivo pelo qual o trabalho foi realizado, e a metodologia utilizada para desenvolver essa dissertação. Os objetivos também são incluídos nessa etapa, o objetivo geral, que traz o foco principal e os objetivos secundários, que são alcançados ao logo do percurso que leva ao objetivo principal.

O segundo capítulo é dividido em duas partes, a primeira apresenta a base de conhecimento necessária para formar a fundamentação teórica, que é feita através das principais referências do SAD FITradeoff e da utilização da neurociência para decisão. A segunda parte, traz uma revisão de literatura, apresentando os principais trabalhos na área de interesse dessa obra. Por fim, mas ainda na segunda parte desse capítulo, é apresentado o posicionamento desse trabalho em relação a outros dessa mesma área, e como ele se diferencia dos demais.

O terceiro capítulo mostra a estrutura do experimento por trás desse trabalho, apresenta os aparatos necessários para aplicação, assim como suas principais particularidades. Outros pontos importantes desse capítulo são a descrição da amostra, assim como a etapa do FITradeoff a qual o estudo é embasado.

O quarto capítulo apresenta os resultados obtidos através dos dados coletados pelo experimento, trazendo de forma conjunta os discursões pertinentes e os pontos analisados.

O quinto capítulo, faz a conclusão das ideias apresentadas e discutidas no quarto capítulo, além de oferecer insights. Nesse capítulo também é oferecido sugestões para trabalhos futuros e pontos de melhorias, discutindo, junto aos dois pontos anteriores, as limitações desse trabalho.

2. REFERENCIAL TEÓRICO E REVISÃO DA LITERATURA

Esse capítulo apresentará um embasamento teórico e uma revisão de literatura, de forma a criar uma fundação sólida dos conhecimentos necessários e das ideias apresentadas nessa obra.

2.1. Referencial Teórico

2.1.1. FITradeoff

A todo momento é possível se deparar com a necessidade de fazer uma escolha ou com o desafio de tomar uma decisão e a importância de escolher a melhor alternativa, dentre todas as possibilidades, aumenta em decorrência da complexidade, dos bônus e ônus que estão atrelados a uma má decisão ou a não tomar a decisão. Assim surge a teoria e as metodologias que tratam da decisão multicritério, uma vez que a melhor escolha nem sempre é óbvia e pode envolver múltiplos objetivos e critérios, que precisam ser analisados simultaneamente. Outro desafio recorrente é a incorporação das preferências do decisor na modelagem do problema, o que torna cada problema de decisão multicritério único e individual. Segundo Saaty (1980), os métodos multicritério ajudam a estruturar e a lidar com a complexidade entre os critérios, atribuindo pesos conforme sua importância relativa para o decisor. Essas abordagens são ainda mais efetivas quando deparadas com problemas que tenham critérios ou objetivos conflitantes, que precisam ser ponderados para fornecer a melhor alternativa.

Os métodos multicritério são basicamente divididos em dois grupos, o primeiro grupo é aquele que admite uma racionalidade não compensatória. Esses métodos não permitem um tradeoff entre os critérios, logo um desempenho inferior não pode ser compensado por um desempenho superior. De acordo com Belton e Stewart (2002), esses métodos utilizam a relação de dominância entre os critérios e as análises lexicográficas, sendo ideais para aplicar em problemas que existam critérios que são considerados indispensáveis. Normalmente essas condições acontecem quando tratamos padrões mínimos que devem ser atendidos ou restrições que não podem ser ultrapassadas.

Por outro lado, o segundo grupo que é o grupo compensatório, enxerga o tradeoff entre os critérios, eles aceitam que o desempenho superior sobressalte uma deficiência em outro critério. Conforme Keeney e Raiffa (1993), esses métodos são apropriados para problemas necessitam de flexibilidade e buscam balancear o desempenho das alternativas através do tradeoff dos critérios. Esses métodos estão diretamente relacionados com as preferências do

decisor e com ele percebe a importância de cada critério, dessa maneira justificando a individualidade e particularidade do processo de decisão.

Independentemente do método utilizando, um grande desafio é encontrado na elicitação das constantes de escala. Para Saaty (1980), as constantes de escala capturam as preferências subjetivas do decisor e permite a combinação dos critérios em uma função normalizada e com escala consistente. Apesar da importância, o processo de elicitação nem sempre é simples de ser realizado, uma vez que é exigido do decisor um grande esforço cognitivo e mental.

Tendo em vista a importância das constantes de escala, das perdas e inconsistências durante o processo de elicitação, De Almeida *et al* (2016) apresentam o FITradeoff, um novo método para elicitação das constantes de escala nos modelos aditivos, abordando uma metodologia de tradeoff flexível e interativo.

O FITradeoff está dentro do escopo da teoria de valor multiatributo (MAVT) e é baseado no processo de Tradeoff tradicional (KEENEY & RAIFFA, 1976), contudo diferente dos métodos clássicos o FITradeoff utiliza o conceito de informações parciais para simplificar o processo de elicitação das constantes de escala, exigindo menos esforço cognitivo por parte do decisor. Como indicado por De Almeida *et al.* (2016), o tradeoff interativo e flexível utiliza o conceito de informações parciais que diz a respeito a dados incompletos e limitados a respeito das preferências do decisor, reduzindo o esforço cognitivo necessário, uma vez que o FITradeoff permite que o decisor forneça apenas informações relevantes.

Como dito anteriormente, o método FITradeoff atua sobre o processo de elicitação das constantes de escala. De Almeida *et al* (2016) descrevem o funcionamento do método através da análise do conjunto de alternativas dentro do espaço de pesos, dentro desse contexto, as alternativas são classificadas em três situações: Ótima, dominada e potencialmente ótima. Dentre os três estados, o estado potencialmente ótimo é crucial, já que à medida que o espaço de pesos é reduzido qualquer alternativa pode se tornar potencialmente ótima. Enquanto existir alternativas potencialmente ótimas, o decisor continuará respondendo perguntas para reduzir cada vez mais o espaço de pesos. Assim, se faz necessário a aplicação de programação linear, com função objetivo definida pela Equação 1.

$$Max Z = \sum_{i=1}^n k_i v_i(x_{ij}), j = 1, 2, \dots, m \quad (1)$$

Em que, x_{ij} é o desempenho da alternativa j no critério i e k é o valor da constante de escala. Contudo, essa função objetivo é submetida a um conjunto de restrições que pode ser visto no conjunto de Equações 2.

$$\begin{aligned}
 \sum_{i=1}^n k_i v_i(x_{ij}) &\geq \sum_{i=1}^n k_i v_i(x_{iz}), z = 1, 2, \dots, m, z \neq j \\
 k_{i+1} &\leq k_i v_i(x_i') - \varepsilon, \text{ para } i = 1 \text{ até } n - 1 \\
 k_{i+1} &\geq k_i v_i(x_i'') + \varepsilon, \text{ para } i = 1 \text{ até } n - 1 \\
 \sum_{i=1}^n k_i &= 1 \\
 k_i &\geq 0, i = 1, 2, \dots, n
 \end{aligned} \tag{2}$$

Em que o conjunto de equações apresenta as restrições de normalização, restrições de não-negatividade, restrições de preferência, restrições de indiferença, restrições de dominância e restrições de comparação das consequências. De maneira geral, o conjunto de restrições tem como objetivo refletir as preferências do decisor, de maneira que a solução final seja garantida e não apenas maximize a função objetivo, mas também respeite as relações de preferência e as condições do decisor. Entregando um resultado robusto e condizente com as expectativas e que respeite as informações e preferências do decisor.

O FITradeoff está longe de ser apenas um método para a problemática de escolha. Frej *et al.* (2019) apresentam a versão do FITradeoff que engloba a problemática de ranking, essa nova abordagem apresenta algumas diferenças quando comparada com o método que aborda apenas a problemática de escolha. A primeira diferença está no estabelecimento das relações de preferência, onde no FITradeoff tradicional o decisor estabelecia indiferença entre as alternativas, já na problemática de ranking o decisor também pode estabelecer relações de preferência estrita. A segunda diferença está na flexibilidade, essa nova abordagem oferece mais flexibilidade ao método, uma vez que o decisor pode interromper o processo de elicitação a qualquer momento que julgar suficiente a quantidade de informações. A terceira diferença está no resultado esperado, a abordagem tradicional de escolhe oferece ao decisor a melhor alternativa, por outro lado na abordagem de ranking o FITradeoff oferece uma ordem completa ou parcial das alternativas a cada interação.

De acordo com Kang *et al.* (2020), a nova abordagem é baseada no processo de elicitação flexível e utiliza limites inferiores e superiores para definir classes consecutivas. A

classificação das alternativas é feita mediante regras de decisão que comparam os valores de borda de cada classe com o valor de comparação de cada alternativa, quando o valor de comparação da alternativa é compreendido entre dois valores de borda a alternativa pertence àquela classe.

Frej *et al* (2021) incluem ao FITradeoff a resolução das problemáticas de portfólio. A abordagem de portfólio tem como objetivo selecionar um conjunto de alternativas que otimize o uso de recursos limitados, dessa maneira maximizando os objetivos em questão.

2.1.2. Neurociência para Decisão

O cérebro humano é uma máquina com infinitas possibilidades, no qual a plasticidade nos permite moldar e redefinir nossas capacidades ao longo da vida. De maneira geral, o cérebro ainda é um mistério e várias questões continuam sem resposta, como a forma que os impulsos elétricos se transformam na personalidade humana, ou ainda como nossa personalidade se mantem mesmo após as células cerebrais se renovarem. Assim, foi necessário criar uma área que estudasse o cérebro, seu funcionamento e seus padrões.

Dessa forma surge a neurociência, que se propõem a estudar o funcionamento do sistema nervoso e os processos cerebrais e tem sido amplamente aplicada para compreender e melhorar o processo de tomada de decisão. Como observou Damasio (1994), as emoções e os aspectos cognitivos desempenham um papel importante na formação de preferências e na avaliação de riscos, influenciando as escolhas realizadas por cada pessoa. Para compreender como esses aspectos influenciam o comportamento dos indivíduos, ferramentas foram criadas para captar informações e métricas referentes aos padrões comportamentais.

Ao longo do tempo várias ferramentas foram desenvolvidas e as que mais se destacam são: eye-tracking, eletroencefalograma (EEG), ressonância magnética funcional (fMRI) e os sensores de resposta galvânica da pele (GRS). Segundo Wedel e Pieters (2008), O eye-tracking mede as métricas relacionadas a movimentação ocular, fornecendo dados sobre o tempo de fixação, dilatação da pupila e número de fixações. As métricas permitem fazer inferências das percepções e do nível de atenção em processos decisórios. O EEG, como discutido por Luck (2014), capta a atividade elétrica cerebral, identificando padrões associados ao processo de decisão em diversas condições, como pressão e incerteza. Raichle e Mintun (2006) discutem a utilização do fMRI como forma de mapear a atividade cerebral em diferentes áreas e diferentes funções cognitivas e comportamentais. Quando exposto ao contexto de decisão, o fMRI destaca as regiões cerebrais ativadas durante a avaliação de uma alternativa. Para Critchley (2002), o

GRS capta como os mecanismos cerebrais processa e responde a estímulos emocionais. Essas ferramentas oferecem uma visão holística sobre os aspectos que influenciam a atividade cerebral e como essa influência é traduzida em métricas captadas por essas ferramentas.

A neurociência e suas ferramentas integrada aos SADs traz como principal vantagem a interpretação dos padrões cognitivos, fisiológicos e neurais, e como essas informações se relacionam com o processo de tomada de decisão. Riedl *et al.* (2010) sugerem que as tecnologias neurocientíficas identificam os padrões subjacentes que influenciam o comportamento do decisor. Oferecendo interpretações mais completas sobre o processo decisório, avaliando fatores racionais, emocionais e aspectos cognitivos subjacentes, que estão ligadas as escolhas feitas ao longo do processo de decisão.

Contudo, a neurociência não está apenas relacionada com a tomada decisão, ela se evidencia em diversas outras áreas, como a neuroeconomia. Essa surge como uma disciplina que combina teorias econômicas, psicologia e neurociência, utilizando para investigar os padrões neurais em aspectos relacionados e economia. Segundo Glimcher e Fehr (2013), essa área tem como objetivo a construção de um modelo mais realista, compreendendo de forma mais profunda os padrões envolvendo decisão sob risco e questões financeiras. Para Kahneman e Tversky (1979), a neuroeconomia revelou que essa preferência é modulada por regiões do cérebro associadas à emoção, como a amígdala e o córtex pré-frontal ventromedial. Estudos como os de De Martino *et al.* (2006) demonstram que, em situações de escolha, os indivíduos tendem a reagir de forma mais intensa a possíveis perdas do que a ganhos de magnitude equivalente, o que pode ser observado através de respostas neurais em áreas específicas do cérebro. Essas descobertas oferecem insights sobre o comportamento do consumidor e sobre como as decisões financeiras podem ser influenciadas por vieses cognitivos e emocionais, muitas vezes desconsiderados em modelos econômicos tradicionais.

Outra aplicação partida da neurociência é o neuromarketing, que busca compreender como o cérebro humano responde a estímulos de marketing e consumo, para investigar como emoções e processos cognitivos influenciam as escolhas do consumidor. A luz de Plassmann *et al.* (2012), o neuromarketing permite identificar preferências cognitivas subjacentes e melhorar estratégias de comunicação e design de produtos. Uma das vantagens do neuromarketing está na capacidade de capturar dados que superam as limitações dos métodos tradicionais de pesquisa de mercado, como questionários. As ferramentas GSR e o EEG permite medir respostas fisiológicas a estímulos, fornecendo insights sobre emoções e engajamento que não são facilmente verbalizados pelos consumidores. Lindstrom (2008) destaca que campanhas baseadas em neuromarketing têm maior probabilidade de criar uma conexão emocional com os

consumidores, aumentando a efetividade das mensagens e a fidelidade à marca. Além disso, o neuromarketing tem sido aplicado para otimizar a disposição de prateleiras, anúncios digitais e interfaces de websites, criando experiências mais intuitivas e agradáveis para os usuários.

A neurociência também pode ser aplicada para avaliar o comportamento dos usuários durante a interação com sistemas de informação, permitindo a compreensão de como elementos de interface afetam a experiência e a usabilidade cuja área é denominada *NeuroIS*. Ferramentas como o eye-tracking são amplamente utilizadas para mapear padrões de atenção e identificar áreas de interesse em interfaces gráficas, revelando quais componentes capturam mais atenção. Conforme apontado por Duchowski, (2007), a medição da atividade cerebral por meio de EEG tem auxiliado a avaliar o nível de esforço cognitivo e frustração do usuário ao interagir com sistemas complexos. As técnicas da neurociência fornecem insights em tempo real, permitindo ajustes que tornam os sistemas mais intuitivos e alinhados às expectativas dos usuários. Outra abordagem relevante é a análise de respostas emocionais dos usuários durante a interação com sistemas, utilizando sensores como o GSR e ferramentas de análise facial. Essas técnicas permitem identificar emoções como frustração, excitação ou satisfação, oferecendo dados objetivos para melhorar a experiência do usuário. Como apontado por Riedl *et al.* (2014), a integração de dados neurocientíficos em sistemas de informação pode ser utilizada para personalizar interfaces e adaptar funcionalidades de acordo com o usuário, promovendo maior engajamento e produtividade. Essas aplicações demonstram como a neurociência pode contribuir para o desenvolvimento de sistemas mais eficazes, capazes de atender às necessidades dos usuários de maneira empática e baseada em dados.

Portanto, ao integrar a neurociência com SAD, é possível não apenas aprimorar o entendimento sobre como decisões são tomadas, mas também desenvolver intervenções mais eficazes. A aplicação dessas ferramentas facilita a inovação em diversas áreas, desde o design de produtos até a elaboração de interfaces mais amigáveis de softwares. Na visão de Riedl *et al.* (2010) o futuro da tomada de decisão está intrinsecamente ligado à capacidade de capturar e interpretar padrões cognitivos por meio de tecnologias neurocientíficas, promovendo escolhas mais conscientes, eficazes e alinhadas aos objetivos organizacionais e individuais.

2.2. Revisão da Literatura

2.2.1. Estudos de Neurociência para Decisão

Essa subseção se propõe a estudar os principais trabalhos que exploraram a neurociência dentro de um contexto de tomada de decisão, explorando as principais ferramentas e suas aplicações.

Ares *et al.* (2014) investigam como os consumidores processam as informações dos rótulos dos produtos. O trabalho identificou dois perfis de consumidores, os analíticos-rationais, que buscam mais informações e analisam dados nutricionais com maior cuidado, e os intuitivo-experienciais, que tomam decisões mais automáticas. Os movimentos oculares dos participantes revelaram diferenças significativas na forma como cada grupo interage com os rótulos.

Bault *et al.* (2016) também utilizam o eye-tracking para investigar os sentimentos de decepção e arrependimento influenciam a tomada de decisão. O trabalho constatou que o arrependimento que foi vivenciado pelo participante está associado a respostas emocionais mais relevantes do que a decepção. A pesquisa foi realizada com base em um experimento de loteria probabilística, onde eram fornecidos feedbacks parciais e completos, sobre o resultado da loteria. Os resultados obtidos pelo rastreamento ocular confirmaram que os participantes compararam seu resultado com o resultado da loteria não escolhida na condição de feedback completo e ignoraram o resultado não realizado da opção escolhida.

Brocas (2012) explora a integração da neurobiologia em modelos de tomada de decisão econômica. São apresentados dois modelos neurocientíficos que descrevem dados e escolhas neurais, e, em seguida, dois modelos econômicos que incorporam essas limitações em uma estrutura baseada na utilidade esperada.

Bechara e Damasio (2005) também abordam as decisões econômicas utilizando aspectos da neurociência. O trabalho sugere que a tomada de decisão racionais dependem de um processamento emocional anterior. O trabalho também levanta a hipótese do marcador somático como forma de oferecer uma estrutura cognitiva para a tomada de decisão, e como esse marcador é influenciado pelas emoções e sentimentos. Apesar de existir a influência, essa nem sempre é consciente, em muitos casos são não conscientes, mas afetam diretamente as decisões econômicas.

Destaca-se ainda o trabalho de Buettner (2016) que utiliza o eye-tracking para investigar a relação entre a complexidade de um site e a carga cognitiva do usuário. Inicialmente foi definido três layouts para o website e os participantes foram expostos a cada um dele e o esforço cognitivo foi registrado através da dilatação da pupila. Os resultados mostraram que o design que menos requer esforço cognitivo é um design moderado utilizando informações textuais e com layout que utilize menus e submenus para navegação dos participantes.

Por fim, a partir da revisão de literatura, constatou-se que existe uma grande quantidade de trabalhos que abordam a neurociência no contexto de decisão, porém cada trabalho discute essa temática dentro dos macros contextos de suas respectivas áreas. As ideias centrais de cada trabalho vão desde estudos voltados para o comportamento durante o processo de escolha até as influências cognitivas nas decisões econômicas, passando de forma contundente sobre o neuromarkting, avaliando o processamento de informações vindo das embalagens dos produtos. Essa grande variedade de temas, indica a abrangência e a flexibilidade que a neurociência e os processos decisório tem sobre diversas áreas do conhecimento. Se mostrando uma ferramenta valiosa na geração de insights e na recomendação de melhorias.

2.2.2. Estudos de Neurociência com o FITradeoff

De forma mais específica que a subseção anterior, esta subseção explora os trabalhos que integraram a neurociência e suas ferramentas com o SAD do FITradeoff. Ao longo da revisão da literatura, houve uma variedade de trabalhos que abordaram esse tema e trataram diferentes aspectos do FITradeoff.

Existe uma parcela considerável de trabalhos que utilizam a neurociência como forma de analisar o comprometimento, engajamento e esforço cognitivo, dos participantes durante o uso do FITradeoff. Os próximos trabalhos citados nessa revisão de literatura, são aqueles mais relevantes dentro do tema de engajamento e esforço cognitivo. De forma mais específica, os trabalhos posteriormente citados, se debruçam sobre a etapa de decomposição do FITradeoff.

Analisando o esforço cognitivo realizado pelo decisor, Roselli *et al.* (2020) investigaram o comportamento dos decisores no processo de elicitação de preferências, através de um experimento utilizando ferramentas da neurociência com um SAD baseado no procedimento Tradeoff no qual os participantes analisaram problemas próprios. Os resultados obtidos através da análise dos dados da pupilometria e do EEG, indicaram um aumento do esforço cognitivo conforme o avanço do experimento, além disso foi constatado baixo engajamento e motivação durante a exploração dos espaços de consequências, assim, inferindo que o Tradeoff é cansativo e favorece inconsistências. Isso se deve a alta demanda cognitiva exigida pelo método.

Da Silva *et al.* (2021) exploraram os aspectos cognitivos do método FITradeoff por meio das ferramentas de eye-tracking e eletroencefalograma. Esta obra investigou a influência que fatores como engajamento e carga cognitiva tem sobre o processo decisório. Através de um experimento, o trabalho indicou que o método FITradeoff facilita a elicitação de preferências com menos perguntas, mostrando vantagens quando comparado a métodos tradicionais. Em

relação aos aspectos citados anteriormente: O trabalho mostrou que o engajamento está relacionado a atenção e a concentração durante o processo de elicitación, já a carga cognitiva explica o esforço mental para processar informações e tomar decisões. Dessa maneira, a pesquisa aponta a importância de considerar ambos no processo de decisão. Foi verificado ainda o impacto da natureza predominante dos critérios do problema de decisão sobre os aspectos cognitivos, sendo os problemas com critérios predominantemente qualitativos demandantes de menos esforço e geradores de maior engajamento.

Da Silva *et al.* (2024) seguiram no mesmo caminho, utilizando eye-tracking e o eletroencefalograma. O trabalho concluiu que além do tipo dos critérios analisadas, a quantidade de critérios envolvidos no problema e a quantidade de perguntas feitas durante a elicitación de preferências influencia diretamente o esforço e o engajamento do decisores. Essa obra também sugeriu aos SAD que estratégias fossem desenvolvidas para minimizar a carga cognitiva sobre o decisor, durante o processo de elicitación de preferências.

Além dos trabalhos que abordam a etapa da decomposição do FITradeoff, há também aqueles que direcionam seu foco para a etapa de avaliação holística. Esses estudos investigam a forma que o decisor expressa suas preferências nesse tipo de avaliação, contribuindo diretamente para a assertividade e confiabilidade da decisão. As próximas obras aqui apresentadas, utilizam ferramentas da neurociência como forma de analisar os processos subjacentes ao longo da avaliação holística.

Roselli *et al.* (2020) apresentaram melhorias para o SAD FITradeoff dentro do contexto da problemática de classificação. Os resultados nesse estudo tiveram base no estudo comportamental utilizando ferramentas de *NeuroIS*. Os dados foram coletados pelo eye-tracking e registraram o comportamento do decisor. A partir dos resultados, foi sugerido aos analistas orientações de suporte ao decisor durante o processo. Acima de tudo, o trabalho evidenciou a importância da etapa de avaliação holística e sugeriu sua inclusão no contexto abordado. Almeida *et al.* (2020) também utilizaram as ferramentas do *NeuroIS* para aprimorar o FITradeoff, contudo diferente do trabalho anterior que utilizou o eye-tracking, esse trabalho atua sobre a visualização holística utilizando o eletroencefalograma.

Roseli *et al.* (2022) abordaram o diagrama alfa-teta como ferramenta para analisar a avaliação holística no processo de decisão. O experimento investigava o esforço cognitivo e o engajamento, quando os participantes avaliam as representações gráficas e tabulares. Os testes mostraram que houve significância estatística no esforço e engajamento durante o experimento.

Ainda falando de trabalho que abordam estudos comportamentais Ferreira *et al.* (2023), abordaram a melhor forma de representação visual na avaliação holística, através de pesquisas

que comparavam gráficos de barra e tabelas. Além de comparar os dois métodos de visualização, o trabalho considerou dois processos de decisão o de escolha da melhor alternativa e o de eliminação da pior alternativa, o que é uma inovação quando comparada com versões anteriores do SAD do FITradeoff. De modo geral, esse trabalho é focado no FITradeoff, porém seus resultados podem ser ampliados para qualquer método que utiliza o MAVT.

De Vasconcelos *et al.* (2024) abordaram os dois modos de avaliação das alternativas, o modo de seleção e o de eliminação, durante a avaliação holística. Esse trabalho se fundamenta na dificuldade de capturar o esforço cognitivo real feito pelos decisores. O trabalho estudou o comportamento dos decisores durante o processo de tomada de decisão, visando modificar o método incluindo aspectos comportamentais. Para realizar as análises, um experimento foi aplicado, onde oito visualizações são apresentadas e os decisores são expostos a modo de seleção e eliminação das alternativas.

Outra parcela dos trabalhos envolvendo neurociência e FITradeoff estuda a melhoria no SAD e a interação com o usuário. Esses trabalhos têm como objetivo adaptar o SAD do FITradeoff a diferentes perfis e contextos. Os próximos trabalhos apresentados abordam as modulações sugeridas.

O trabalho de Roselli *et al.* (2022) utilizou estudos comportamentais de neurociência para modular o método FITradeoff. Esse trabalho sumarizou todas as modulações sofridas pelo FITradeoff ao longo do tempo, as modulações foram conduzidas de duas maneiras, a primeira é a modulação no processo de elicitación de preferências e a segunda foi realizada no software de apoio a decisão do FITradeoff. A primeira modulação ofereceu insights e recomendações aos analistas para orientar os decisores durante o uso do método, já a segunda modulação trouxe mudanças no sistema deixando mais eficiente e assertivo.

De Almeida *et al.* (2023) apresentaram uma visualização completa do método FITradeoff. Esse trabalho abordou todas as aplicações oferecidas pelo FITradeoff, bem como o desenvolvimento metodológico partindo da integração do método com ferramentas da neurociência e de experimentos comportamentais. De forma geral, os autores propuseram explorar a eficácia das avaliações holísticas em reduzir o número total de perguntas que o decisor precisa responder, além de encorajar a aplicação prática do método nas mais diversas organizações, fortalecendo assim a conexão entre desenvolvimentos metodológicos e aplicações no mundo real. O trabalho de De Almeida *et al.* (2024) segue a mesma linha, apontando a possibilidade de alternar entre as duas metodologias para melhorar a consistência das preferências expressas, além de deixar o processo de elicitación mais fluido, focando

contrudo, na estrutura axiomática do método após diversas modulações decorrentes de estudos de estudos comportamentais e de neurociência.

Por fim, é possível concluir que existem uma gama de trabalhos que relacionam a tomada de decisão com as ferramentas de neurociência, em sua maioria os trabalhos se utilizam das ferramentas de eye-tracking e eletroencefalograma. A preferência por essas duas ferramentas está relacionada com a natureza do SAD FITradeoff combinado com os experimentos que são aplicados, uma vez que o sistema tem representações visuais relevantes, além de aspectos ligados a preferência cognitiva subjacente dos participantes. Assim, justificando o grande número de trabalhos que se utilizam essas duas ferramentas.

2.2.3. Síntese do Estado da Arte e Posicionamento deste Trabalho

Partindo da revisão de literatura realizada nessa obra, verificou-se que há um número significativo de trabalhos na área de apoio a decisão com neurociência. Essas obras abordam contextos variados como decisões econômicas e de compra, por outro lado, os trabalhos que se dedicam a análise de decisão multicritério são mais limitados, porém significativos para o método FITradeoff.

Parte dos trabalhos sobre decisão multicritério, tem focado na avaliação intercritério, esses estudos se concentram nos dois paradigmas que são a elicitación por decomposição e avaliação holística. Esses estudos utilizam ferramentas da neurociência como eye-tracking e EEG ou estudos puramente comportamentais. Essa dissertação se diferencia, por sua vez, por se debruçar na avaliação intracritério por meio do método da bisseção, conforme incorporado no SAD do FITradeoff. Esta obra examina como o gênero e o nível de conhecimento influenciam aspectos cognitivos, como atenção e aprendizado. Com a finalidade de fornecer insights que permitam, aos desenvolvedores e analistas de decisão, melhorias para o processo decisório como um todo. Além disso, este trabalho busca levantar insights que sirvam para outros métodos e SADs relacionados.

3. EXPERIMENTO

A aplicação do experimento, proposto nesse trabalho, foi composta de nove etapas, como ilustra a Figura 2. Cada etapa foi pensada e desenvolvida para que os resultados alcançados fossem o mais fidedigno possível. O experimento foi aplicado com estudantes dos cursos de graduação e pós-graduação em engenharia de Produção da Universidade Federal de Pernambuco e profissionais atuantes no mercado de trabalho.

Figura 2 - Etapas do experimento



Fonte: O autor (2024).

A infraestrutura do experimento foi dividida em duas salas, na primeira ocorria a etapa de recepção dos participantes onde eles aguardavam sua vez, os agendamentos eram feitos com horários pré-estabelecidos para que não houvesse um tempo de espera demasiado e nem alterações no humor que pudessem influenciar no resultado do experimento.

Na segunda etapa ocorreu a explicação das ferramentas de neurociência que iriam ser aplicadas, de forma conjunta, foi explicado acerca dos procedimentos do experimento. Em particular, sobre o uso do EEG, que necessita aplicar um gel abrasivo, que é aplicado no couro cabeludo para melhorar as conexões e os sinais do aparelho. Após a explicação sobre o gel, também foi explicado aos participantes que ao fim do experimento seria disponibilizado produtos para limpeza e higienização, casos achassem necessário.

Após a explicação na segunda etapa e com a aceitação do participante em realizar o experimento, é pedido que um formulário seja preenchido juntamente com a assinatura do Termo de Consentimento Livre e Esclarecido (TCLE) do Comitê de Ética da UFPE (CAAE - 69253017.0.0000.5208). Esse documento estabelece o consentimento do participante e esclarece, de forma clara e objetiva, todas as informações a respeito do experimento e da pesquisa.

A quarta etapa, é o posicionamento do eletroencefalograma que tem como componente uma touca com 24 canais de transmissão, que são conectados a um aparelho que capta as ondas cerebrais e as convertem de modo que possam ser lidas e analisadas. Nessa etapa se faz necessário a utilização do gel abrasivo no couro cabeludo, esse gel tem a finalidade de diminuir

a interferência causadas pela sujeira, oleosidade do couro cabeludo, quantidade de cabelo e eventuais interferências que possam surgir entre o conector e a cabeça do participante. Apesar do nome sugestivo, esse produto não traz nenhum tipo de desconforto ao ser aplicado e é retirado facilmente, pelo participante, ao lavar o cabelo normalmente.

Figura 3 - Aparelhos de eletroencefalograma utilizado no experimento



Fonte: O autor (2024).

Durante a quinta etapa ocorre a calibração do Eye-Tracking, o aparelho utilizado nesse experimento segue as seguintes especificações: Eye-tracking 120 da Tobii que estava conectado a um notebook de 24 bits, com sistema operacional Windows 10 de 4GB de memória RAM e um monitor com resolução de 1280 x 1024 pixels, além de acessórios básicos como teclado e mouse. Devido a sensibilidade do rastreador ocular, 120Hz de taxa de amostragem, 0,5° de precisão e tem capacidade de captar olhares de até 35° de angulação, foi preciso estabelecer medidas padrões para que o experimento fosse realizado sempre da mesma forma, independente do biotipo do participante. Os parâmetros estabelecidos foram:

- Distância do monitor: A tela foi posicionada a 60 cm dos participantes;
- Angulação do Eye-tracking: A inclinação entre a mesa onde o aparelho estava apoiado e o coletor era aproximadamente 30°;
- A distância entre Eye-tracking e a tela: O aparelho foi posicionada a 5cm da base do monitor;
- Altura da tela: A altura do monitor em relação a mesa foi de 15 cm;

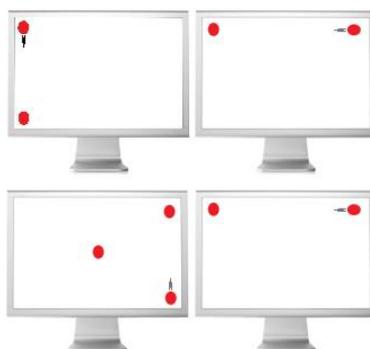
Outra atividade importante na etapa 5 é a calibração do aparelho de rastreamento ocular. Esse ajuste ocorre em duas etapas, a primeira etapa é a da cruz de ficção, mostrada na Figura 4 seguida pela calibração por perseguição de pontos aleatórios, onde uma bolinha vermelha se movimenta de forma aleatória e o participante a segue apenas com os olhos, evitando os movimentos complementares de cabeça e tronco.

Figura 4 - Cruz de fixação para calibração do Eye-tracking



Fonte: Autor (2024).

Figura 5 - Calibração por perseguição de pontos aleatórios



Fonte: O autor (2024).

Após as ferramentas calibradas e ajustadas, o participante inicia a sexta etapa, o uso do método FITradeoff. Essa etapa se inicia com o participante efetuando login na página do método, como apresentado na Figura 6, e, de forma sequencial, o usuário deve escolher que tipo de problemática é envolvida (ver Figura 7). A partir de então, segue-se a elicitacão de preferências.

Após o resultado do FITradeoff, os aparatos são tirados do participante e são oferecidos produtos de limpeza e higiene, para retirada do gel abrasivo, nessa etapa o participante pode escolher se deseja remover o gel na hora ou posteriormente em um lugar diferente. Por fim, o participante responde um segundo questionário, que busca coletar informações do seu perfil socioeconômico, em seguida ele é liberado para sair da sala de experimento.

Figura 6 - Tela de Login da página do FITradeoff

FITradeoff.org
FITradeoff
 Flexible and interactive Tradeoff

FU-T3MMM-WF1a

Help

E-mail:

Password:

[Forgot password?](#)

Back Enter

inct INSID CDSID UFPE

Fonte: SAD FITradeoff (2024).

Figura 7 - Escolha da problemática

FITradeoff.org
FITradeoff
 Flexible and interactive Tradeoff

FU-T3MMM-WF1a

Logout

Please, select an image to choose a module and press continue:

Choice

Ranking

Portfolio with Benefit-to-Cost

Sorting

Continue

inct INSID CDSID UFPE

Fonte: SAD FITradeoff (2024).

3.1. Problema de decisão

Para a realização desse experimento, foi apresentado um problema de escolha que tinha como propósito a compra de um smartphone. O problema foi modelado através de sete critérios e sete alternativas, como mostrado na Tabela 1. As alternativas representam diferentes smartphones enquanto os critérios são os requisitos técnicos, os participantes, por sua vez, deverão julgar cada alternativa por meio dos seus respectivos desempenhos em cada um dos critérios utilizando o SAD do FITradeoff.

De forma paralela e integrada, um aparelho de rastreamento ocular coleta as informações a respeito dos olhos, como métricas relacionadas à dilatação da pupila e número de fixações, enquanto os participantes utilizam o FITradeoff para decidir qual a melhor opção de smartphone com base nas suas preferências.

Tabela 1 - Alternativas e critérios

Alternativas	Critérios						
	Preço (R\$)	Qualidade e das fotos (MP)	Armazenamento (GB)	Bateria (Horas)	Tela (Polegadas)	Design	Ano de lançamento
A1	9.269,00	5	512	25	6,7	5	2023
A2	6.999,00	5	128	20	6,1	5	2022
A3	8.549,00	5	256	25	6,8	5	2023
A4	1.349,00	3	128	18	6,6	2	2021
A5	4.549,00	4	128	18	6,1	4	2022
A6	2.199,00	3	128	20	6,5	3	2023
A7	942,00	1	32	18	6,2	1	2020

Fonte: O autor (2024).

3.2. Amostras

Os dados que compõem essa pesquisa são oriundos de uma amostra que é composta por quatro grupos, os grupos são detalhados na Tabela 2. Dois fatores diferenciam os grupos, o gênero e o nível de conhecimento sobre decisão multicritério (DM), de forma geral todos os participantes eram alunos da graduação ou pós-graduação e tinham ou não cursado a disciplina de DM, portanto definindo se os participantes tinham ou não conhecimento acerca de DM.

Apesar da existência de quatro grupos, cada um com suas devidas características, há um aspecto comum entre eles, todos os participantes envolvidos no experimento estavam ativos no mercado de trabalho e ocupavam posições dos decisores. Ainda comentando sobre a amostra, a idade dos participantes varia em uma faixa etária de 22 a 41 anos de idade.

Tabela 2 – Grupos que compõem a amostra

Grupo	Quantidade
Homens com conhecimento	15 participantes
Mulheres com conhecimento	8 participantes
Homens sem conhecimento	13 participantes
Mulheres sem conhecimento	6 participantes

Fonte: O autor (2024).

3.3. Avaliação intracritério

É na etapa de avaliação intracritério que as preferências do decisor são quantificadas, essa etapa se propõe a analisar de forma isolada e detalhada como cada alternativa se comporta em um único critério, possibilitando à comparação e posteriormente a determinação da melhor alternativa. Segundo Almeida (2013), a função valor que relaciona a alternativa i com o critério j é dada por $v_j(a_i)$, ela expressa o desempenho das alternativas em uma escala de 0 a 1, onde 0 é o menor desempenho e 1 é o desempenho máximo. Por outro lado, a função valor também permite representar o problema de decisão através de uma forma matricial, como pode ser visto na Tabela 3, essa matriz é amplamente conhecida como matriz de decisão e a forma de input adotada por diversos métodos de decisão multicritério.

Tabela 3 - Matriz de decisão

Alternativas	Critérios			
	C_1	C_2	...	C_m
a_1	$v_1(a_1)$	$v_2(a_1)$...	$v_m(a_1)$
a_2	$v_1(a_2)$	$v_2(a_2)$...	$v_m(a_2)$
...
a_n	$v_1(a_n)$	$v_2(a_n)$...	$v_m(a_n)$

Fonte: Adaptado de Almeida (2013).

O SAD FITradeoff acompanha o decisor ao longo de todas as etapas do processo decisório, desde o input dos dados de entrada, através do Excel ou diretamente no SAD, até o output e análise de sensibilidade. A Figura 8 mostra de forma mais específica a interface do FITradeoff ao longo da etapa de avaliação intracritério.

Figura 8 - Interface do FITradeoff na etapa intrecritério

FITradeoff.org
FITradeoff
Flexible and Interactive Tradeoff

FU-T30004-WF1a
Logout

Intra-Criteria Evaluation

Help Reset

Select one criterion to start the elicitation: Price (R\$)

You can either:
[Declare as a linear function](#) if you do not want to elicit the marginal value function ⓘ
 Or Proceed to elicitation below.

Considering the criterion "Price (R\$)" comparing the two ranges below,
 which range do you prefer to decrease, in order to have a greater increase in the value of the range?

From 9269 to 5106 .
 From 5106 to 942 .
 Indifferent.

Ok

Fonte: SAD FITradeoff (2024).

Nessa etapa o decisor escolhe qual critério deseja avaliar, a seta verde na Figura 9 indica qual critério está sendo avaliado naquele instante, ao fim de cada avaliação, o decisor escolhe o próximo critério até que todos sejam avaliados e o FITradeoff passe para a próxima etapa.

Figura 9 - Seleção de critérios na avaliação intracritério

FITradeoff.org
FITradeoff
Flexible and Interactive Tradeoff

FU-T30004-WF1a
Logout

Intra-Criteria Evaluation

Help Reset

Select one criterion to start the elicitation: Price (R\$)

You can either:
[Declare as a linear function](#) if you do not want to elicit the marginal value function ⓘ
 Or Proceed to elicitation below.

Considering the criterion "Price (R\$)" comparing the two ranges below,
 which range do you prefer to decrease, in order to have a greater increase in the value of the range?

From 9269 to 5106 .
 From 5106 to 942 .
 Indifferent.

Ok

Fonte: SAD FITradeoff (2024).

A Figura 10 mostra onde e como é feito a avaliação dos critérios no SAD do FITradeoff. O decisor é questionado a respeito dos ranges de valores em que o critério é elicitado, novos

ranges são gerados até que, por fim, o decisor se mostre indiferente ou a condição de parada seja atendida, e um novo critério precise ser elicitado.

Figura 10 - Avaliação dos critérios na interface FITradeoff

The screenshot displays the FITradeoff web interface. At the top, it says 'FITradeoff Flexible and Interactive Tradeoff'. The main heading is 'Intra-Criteria Evaluation'. Below this, there's a dropdown menu set to 'Price (R\$)'. The text asks the user to either 'Declare as a linear function' or 'Proceed to elicitation below'. A slider is shown with values 9269, 5106, and 942. A blue arrow points from 9269 towards 942. To the right of the slider are three radio button options: 'From 9269 to 5106 .', 'From 5106 to 942 .', and 'Indifferent.'. A green arrow points to the 'Indifferent.' option. There are 'Help' and 'Reset' buttons in the top right corner.

Fonte: SAD FITradeoff (2024).

Em alguns casos o decisor pode reconhecer que existe uma relação linear no desempenho desse critério e pode declarar como uma função linear. Após todas as elicitaciones dos critérios o gráfico da função de valor marginal de cada um é traçado, a função tem como base os pontos de indiferença que foi elicitado pelo decisor.

O papel principal da função de valor marginal é captar e especificar o relacionamento entre as variações no desempenho das alternativas e o impacto que isso causa na preferência do decisor. Almeida (2013) sugere que a etapa intracritério tem uma relação direta com os atributos envolvidos no problema e as escalas utilizadas na avaliação.

Como dito anteriormente, na elicitación intracritério o decisor avalia o ranges, esse método de avaliação é conhecido como bisseção. Nesse método o decisor deve encontrar um ponto na escala que ele perceba valor igualitário entre as extremidades, obtendo para cada escolha funções valor, esse processo pode ser repetido até que o decisor se sinta indiferente entre os valores.

Figura 11 - Declaração de valor linear no SAD FITradeoff

FITradeoff.org
FITradeoff
Flexible and Interactive Tradeoff

PA-TDMH-WFL
Logout

Intra-Criteria Evaluation
Help Reset

Select one criterion to start the elicitation: Price (R\$)

You can either:
[Declare as a linear function](#) if you do not want to elicit the marginal value function ⓘ 
 Or Proceed to elicitation below.

Considering the criterion "Price (R\$)" comparing the two ranges below, which range do you prefer to decrease, in order to have a greater increase in the value of the range?

From 9269 to 5106 .
 From 5106 to 942 .
 Indifferent.

Ok

Fonte: SAD FITradeoff (2024).

3.4. Eye-Tracking

A combinação do eye-tracking e dos SADs tem uma grande importância prática, já que ele detalha o comportamento e as preferências dos usuários através do rastreamento ocular, quando são expostos a interação com sistemas computacionais. A integração dessas metodologias ajuda a entender como os usuários processam as informações e possibilita o ajuste do sistema, deixando a interface mais intuitiva e personalizada.

O experimento foi aplicado utilizando o EEG e o eye-tracking como formas de coletar os dados fisiológicos e cognitivos dos participantes, essa informação pode descrever as preferências cognitivas subjacentes. Apesar de utilizar o eletroencefalograma, essa pesquisa se debruça apenas sobre as métricas captadas pelo eye-tracking, já que estão mais alinhadas com as análises que esse trabalho se propõe a fazer. E por si só já fornece dados valiosos sobre a atenção dos participantes.

Pan *et. al.* (2011) apontam a necessidade de subdividir a tela em áreas menores, de modo a captar com exatidão o foco da atenção dos participantes, tendo isso em vista a Figura 12 traz a subdivisão dos elementos da interface, na etapa intracritério, em áreas de interesse pertinentes ao estudo, já a Tabela 4 descreve cada área de interesse.

Figura 12 - Áreas de interesse

Fonte: O autor (2024).

Tabela 4 - Descrição das áreas de interesse

Área de interesse	Descrição
a	Instruções B
b	Instruções D
c	Gráfico
d	Alternativas

Fonte: O autor (2024).

O eye-tracking é capaz de capturar diversas métricas relacionadas aos olhos, fornecendo insights valiosos acerca do comportamento visual, cognitivo e fisiológico dos participantes. Rayner (1998) descreve os movimentos oculares e alguns deles são listados abaixo:

- **Sacada:** Movimentos rápidos e curtos entre dois pontos de fixação, normalmente são responsáveis por trocar o foco e a atenção visual;
- **Fixação:** Tempo em que o olho permanece parado por um período, em um ponto específico do espaço, durante o período de fixação o cérebro processa as informações visuais;
- **Micro movimentos oculares:** São pequenos movimentos involuntários que ocorre até mesmo quando o olho está fixo em um ponto, eles contribuem para a clareza da visão;
- **Movimento de acomodação:** Ajuste do cristalino ao olho focando objetos;

- Anti-sacada: Movimento realizado no sentido oposto ao ponto de fixação subitamente apresentado;
- Perseguição lenta: São os movimentos de perseguição que tem como finalidade manter a estabilidade da imagem, na fóvea, de objetos em movimento;

Além dos movimentos sugeridos por Rayner (1998), existem outras métricas captadas pelo eye-tracking como, fixações, dilatação da pupila, número de piscadas etc. Esses dados possibilitam uma análise diferenciada quando comparadas apenas aos dados do movimento ocular. A integração desses dois conjuntos de dados oferece contribuições preciosas para os estudos comportamentais, analisando aspectos diferentes e complementares que são traduzidos em assertividade.

Essa pesquisa se delimita a usar duas métricas, o número de fixações e a duração da fixação, apesar de existirem outras tantas métricas que podem agregar a pesquisa, as duas utilizadas responderam bem ao escopo e aos objetivos. Principalmente pela abordagem das áreas de interesse, que necessita de indicadores condizentes para que as análises feitas atendam as expectativas, uma vez que o tempo gasto olhando para uma área de interesse juntamente com o número de vezes que o participante fixou sua atenção em um determinado lugar, revela a condição do participante (estado de dúvida, dispersão ou interesse) ao longo do experimento. *Insights* como esses serão utilizados para fornecer melhorias para o SAD e para contribuir de forma significativa com os futuros analistas que aplicarão o método FITradeoff com decisores que tenham características iguais aos apresentados nessa obra.

3.5. Estatísticas

Além dos dados relacionados com as métricas oculares, esse trabalho também coletou as informações sobre a quantidade de tempo que cada participante levou para finalizar a etapa de avaliação intracritério. O intuito foi apresentar uma análise mais abrangente e que possa detalhar não só as relações que as características dos participantes têm com as áreas de interesse, mas também como essas estão ligadas ao aprendizado ao longo do experimento.

Ao analisar as informações referentes ao tempo de duração da etapa de avaliação intracritério, é possível apresentar insights sobre aspectos cognitivos dos participantes, assim como também é possível comparar, as informações da amostra como um todo, com as informações oriundas de cada um dos grupos descritos na Tabela 2. Dessa maneira, é possível

verificar a existência de diferenças significativas entre os tempos de execução de cada grupo e a relação de entre as características e a taxa de aprendizagem.

Posteriormente às análises da curva de aprendizagem, a pesquisa se envereda através das métricas oculares por meio da análise de regressão. Os estudos utilizando as regressões fornecem dados significativos de como o gênero (masculino e feminino) e o conhecimento (ter ou não ter) influenciam no foco da atenção em áreas de interesse específicas.

As próximas subseções abordarão de forma mais profunda as análises mencionadas anteriormente, apontando as características e a natureza das metodologias usadas nessa pesquisa.

4. RESULTADOS E DISCUSSÕES

Nessa sessão será discutido o desfecho das análises propostas na sessão anterior, assim como a explanação dos testes e ferramentas utilizadas para alcançar as interpretações e resultados.

4.1. Curva de Aprendizagem

Ao realizar qualquer tipo de tarefa, com o passar do tempo ou com o aumento do número da prática, ganha-se experiência e proficiência na execução dessa atividade e isso é revertido na eficiência com que a desempenhamos. A curva de aprendizagem descrita por Wright (1936) descreve o desempenho que uma pessoa tem ao realizar uma atividade, Para Hermann Ebbinghaus (1985) a curva de aprendizagem revela a relação entre o esforço investido e a quantidade de informação retida ao longo do tempo.

Para construir uma curva de aprendizagem relacionada a etapa intracritério, do experimento proposto nesse trabalho, foi preciso coletar os tempos que cada participante levou para concluir cada uma das elicitções das funções de valor intracritério, dos setes critérios envolvidos no problema. A Tabela 5 mostra o tempo que cada participante levou para cada elicitção, assim como o tempo total da etapa de elicitção intracritério. A primeira coluna da Tabela 5 mostra o número dos participantes, porém não de forma sequencial, já que durante a aplicação do experimento os participantes que tiveram problemas na leitura ou na captação dos dados, tiveram que ser excluídos da amostra, a fim de evitar viés e imprecisão nas análises.

A Equação I, descreve a curva de aprendizagem abordada nesse trabalho, em que, o parâmetro b descreve a taxa de aprendizado, com valores que varia entre 0 e -1, quanto mais próximo de -1 maior é a taxa de aprendizado do participante.

$$y = C_1 x^b \quad (I)$$

Quanto ao parâmetro y , é o desempenho do participante, C_1 é o desempenho do participante na primeira execução e x indica a prática do participante.

Tabela 5 – Duração, em segundos, da elicitação de cada critérios, para todos os participantes

	Rodada 1	Rodada 2	Rodada 3	Rodada 4	Rodada 5	Rodada 6	Rodada 7	Total
Participante 01	221,51	66,45	42,87	14,5	11,91	18,38	9,74	385,36
Participante 03	67,3	25,75	32,13	46,43	31,64	25,23	43,46	271,94
Participante 04	61,45	41,08	73,1	48,24	48,1	25,16	36,68	333,81
Participante 05	111,56	45,97	43,17	39,65	13,48	54,03	37,89	345,75
Participante 06	164,86	10,88	75,73	77,67	36,75	21,98	65,9	453,77
Participante 07	144,2	38,81	83,56	96,66	16,66	27,93	27,69	435,51
Participante 08	32,77	28,06	26,77	49,32	20,98	21,82	5,41	185,13
Participante 09	97,57	31,87	20,61	71,93	47,33	24,59	29,18	323,08
Participante 10	64,05	45,44	26,87	35,95	23,1	17,58	41,73	254,72
Participante 11	59,76	17,62	29,77	17,77	25,37	7,88	16,22	174,39
Participante 12	90,79	42,61	42,22	61,07	55,01	21,05	46,18	358,93
Participante 13	84,91	28,46	26,41	28,99	26,44	4,42	21	220,63
Participante 14	56,24	25,12	44,99	40,44	45,37	17,18	50,36	279,7
Participante 15	89,33	43,3	62,98	47,07	79,61	27,98	56,5	406,77
Participante 16	51,25	45,67	44,93	34,36	28,49	18,16	24,58	247,44
Participante 17	82,98	34,25	29,27	45,85	33,26	17,78	21,58	264,97
Participante 18	63,57	64,4	76,29	55,69	45,14	17,89	52,11	375,09
Participante 20	33,47	77,7	34,7	12,29	29,82	22,88	27	237,86
Participante 22	66,37	16,96	34,07	22,91	22,26	33,16	42,88	238,61
Participante 23	145,31	17,17	35,23	25,93	26,71	21,68	77,98	350,01
Participante 24	34,88	34,89	14,99	37,29	44,44	10,33	41,66	218,48
Participante 25	82,01	70,5	23,75	11,66	29	25,07	18,16	260,15
Participante 26	64,83	73,7	32,09	44,72	30,29	13,83	35,73	295,19
Participante 27	73,11	20,27	24,72	19,36	26,39	7,57	22,37	193,79
Participante 28	45,04	13,86	20,13	36,07	35,34	8,09	22,15	180,68
Participante 29	133,12	56,03	66,19	55	101,89	100,71	63,58	576,52
Participante 30	59,52	10,88	20,4	26,63	22,07	9,96	17,8	167,26
Participante 31	160,7	85,66	80,94	26,8	89,23	18,44	36,3	498,07
Participante 32	78,21	22,86	15,2	7,74	23,68	30,81	39,49	217,99
Participante 33	105,62	49,1	46,06	48,1	43,62	27,97	38,28	358,75
Participante 34	55,84	21,01	38,66	34,55	34,16	8,48	21,28	213,98
Participante 35	36,18	18,84	40,44	42,42	32,94	20,35	34,4	225,57
Participante 36	89,11	84,18	22,95	18,73	26,8	13,59	25,41	280,77
Participante 37	67,17	46,22	47,06	49,19	35,77	18,47	42,49	306,37
Participante 38	68,11	54,75	41,54	15,04	37,19	14,39	32,66	263,68
Participante 39	38,2	12,94	22,26	20,91	17,35	11,53	10,46	133,65
Participante 40	79,79	55,19	51,23	37,23	55,05	19,26	31,35	329,1
Participante 41	10,62	4,02	16,46	9,99	7,23	3,07	9,89	61,28
Participante 42	108,48	40,06	63,34	64,96	52,12	19,55	30,8	379,31
Participante 43	95,04	53,79	59,63	86,45	74,59	29,79	60,45	459,74
Participante 44	24,85	29,23	21,74	19,09	21,56	11,67	23,56	151,7
Participante 45	15,9	23,33	12,21	25,77	20,76	5,21	14,34	117,52

Fonte: O autor (2024).

Dentre os inúmeros modelos de curva de aprendizagem existentes na literatura, as curvas que melhor descrevem o aprendizado em tarefas cognitivas são os modelos potenciais. Essas famílias de curvas são as que melhor modelam melhorias progressivas, mas com um ritmo decrescente ao longo do tempo. Os modelos potenciais refletem o fato de que ganhos iniciais, no aprendizado, tendem a ser maiores com avanços subsequentes mais lentos, ou seja, características típicas de tarefas que requerem processamento mental, retenção de informação e resolução de problemas, como dito por Newell e Rosenbloom (1981).

4.1.1. Estatísticas descritiva

Uma análise das estatísticas descritivas tem como principais objetivos: Organizar, entender, resumir e apresentar de maneira clara e objetiva as características de um conjunto de dados. Dito isto, podemos concluir que as estatísticas descritivas oferecem uma base sólida para a interpretação e tomada de decisão, principalmente pela forma sintetizada e visual que as informações são apresentadas. Essa sessão iniciará trazendo as estatísticas descritivas da amostra, de forma subsequente trará as estatísticas descritivas de cada grupo apresentado na Tabela 2, no formato de subseções. Por fim, finalizando essa sessão, será apresentada a comparação entre os grupos e a amostra.

O Tabela 6 apresenta as estatísticas que descrevem o tempo de execução das rodadas, da elicitación dos critérios, da amostra dos participantes, nela é possível ver a média (mean), o desvio padrão (std), o valor mínimo (min) e máximo (max), o primeiro quartil (25%), o segundo quartil (50%) e o terceiro quartil (75%).

Tabela 6 - Estatísticas descritivas da Amostra.

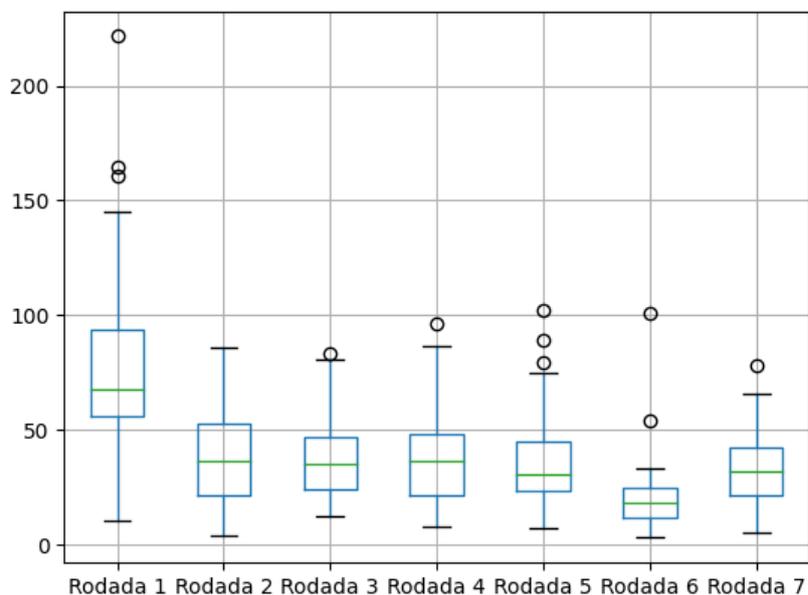
	Rodada 1	Rodada 2	Rodada 3	Rodada 4	Rodada 5	Rodada 6	Rodada 7
count	42.000000	42.000000	42.000000	42.000000	42.000000	42.000000	42.000000
mean	78.942381	38.782857	39.706190	38.343333	36.402381	20.830952	33.492381
std	43.086365	21.276934	19.639052	20.908117	20.247709	15.729651	16.472156
min	10.620000	4.020000	12.210000	7.740000	7.230000	3.070000	5.410000
25%	55.940000	21.472500	23.992500	21.410000	23.245000	12.150000	21.722500
50%	67.705000	36.850000	34.965000	36.650000	30.965000	18.455000	32.005000
75%	93.977500	52.617500	46.810000	48.205000	44.965000	24.950000	42.300000
max	221.510000	85.660000	83.560000	96.660000	101.890000	100.710000	77.980000

Fonte: O autor (2024).

A primeira rodada apresenta uma média de 78,94, com amplitude de 210,89 (diferença entre o máximo de 221,51 e o mínimo de 10,62) e desvio padrão de 43,09, resultando em um coeficiente de variação de 54,58%, indicando uma dispersão moderada relativa à média. Os quartis mostram que 25% dos tempos estão abaixo de 55,94, 50% abaixo de 67,71 (mediana),

e 75% abaixo de 93,98. A Figura 13, que apresenta o boxplot da rodada, evidencia a presença de outliers que impactam a média e a amplitude.

Figura 13 - Diagrama Boxplot para a amostra



Fonte: O autor (2024).

Na segunda rodada, a média é de 38,78, com amplitude de 81,64 e desvio padrão de 21,28, resultando em um coeficiente de variação de 54,87%, semelhante ao da rodada 1. Os quartis indicam que 25% dos tempos estão abaixo de 21,47, 50% abaixo de 36,85 (mediana), e 75% abaixo de 52,62. Apesar da dispersão ser significativa, o boxplot sugere que os valores estão mais centralizados, com menos influência de outliers.

A terceira rodada apresenta uma média de 39,71, amplitude de 71,35 e desvio padrão de 19,64, resultando em um coeficiente de variação de 49,45%, indicando maior controle na dispersão dos dados em relação às rodadas anteriores. Os quartis mostram que 25% dos tempos estão abaixo de 23,99, 50% abaixo de 34,97, e 75% abaixo de 46,81. A Figura 13 confirma que os dados estão concentrados, sem evidência marcante de valores extremos.

A quarta rodada tem uma média de 38,34, com amplitude de 88,92 e desvio padrão de 20,91, resultando em um coeficiente de variação de 54,54%. Os quartis mostram que 25% dos dados estão abaixo de 21,41, 50% abaixo de 36,65, e 75% abaixo de 48,21. Apesar da ampla amplitude, a dispersão relativa está alinhada às rodadas anteriores, e o boxplot sugere que a variabilidade decorre da distribuição natural dos tempos.

A quinta rodada possui uma média de 36,40, amplitude de 94,66, e desvio padrão de 20,25, com um coeficiente de variação de 55,62%, novamente indicando uma dispersão moderada. Os quartis indicam que 25% dos tempos estão abaixo de 23,25, 50% abaixo de 30,97, e 75% abaixo de 44,97. O boxplot da Figura 13 revela alguns valores extremos que contribuem para a alta amplitude.

A sexta rodada apresenta a menor média entre todas, 20,83, com amplitude de 97,64 e desvio padrão de 15,73, resultando em um coeficiente de variação de 75,53%, o mais alto entre as rodadas. Os quartis mostram que 25% dos tempos estão abaixo de 12,15, 50% abaixo de 18,46, e 75% abaixo de 24,95. A Figura 13 confirma a presença de outliers significativos que impactam tanto a amplitude quanto o desvio padrão.

Por fim, a sétima rodada tem uma média de 33,49, amplitude de 72,57, e desvio padrão de 16,47, com um coeficiente de variação de 49,15%, indicando a menor dispersão relativa entre todas as rodadas. Os quartis mostram que 25% dos tempos estão abaixo de 21,72, 50% abaixo de 32,01, e 75% abaixo de 42,30. O boxplot evidencia uma distribuição controlada, com poucos ou nenhum outlier significativo.

A análise geral das rodadas revela que a primeira rodada possui a maior amplitude e os outliers mais extremos, enquanto a sexta rodada apresenta o coeficiente de variação mais alto, indicando maior variabilidade relativa. Por outro lado, a terceira e sétima rodadas apresentam maior controle na dispersão, com coeficientes de variação mais baixos e distribuições mais homogêneas. A presença de outliers na primeira rodada, na quinta rodada e sexta rodada, ressalta a necessidade de uma análise cuidadosa para entender as causas desses valores extremos e como eles podem impactar as conclusões sobre os tempos médios.

4.1.1.1. Estatística descritiva do grupo homens com conhecimento

A Tabela 2 descreve os grupos que compõem a amostra, assim como, a quantidade de participantes que fazem parte desses grupos. O Tabela 7 apresenta as estatísticas que descrevem o tempo de execução das rodadas, do grupo homens com conhecimento, nela é possível ver a média (mean), o desvio padrão (std), o valor mínimo (min) e máximo (max), o primeiro quartil (25%), o segundo quartil (50%) e o terceiro quartil (75%).

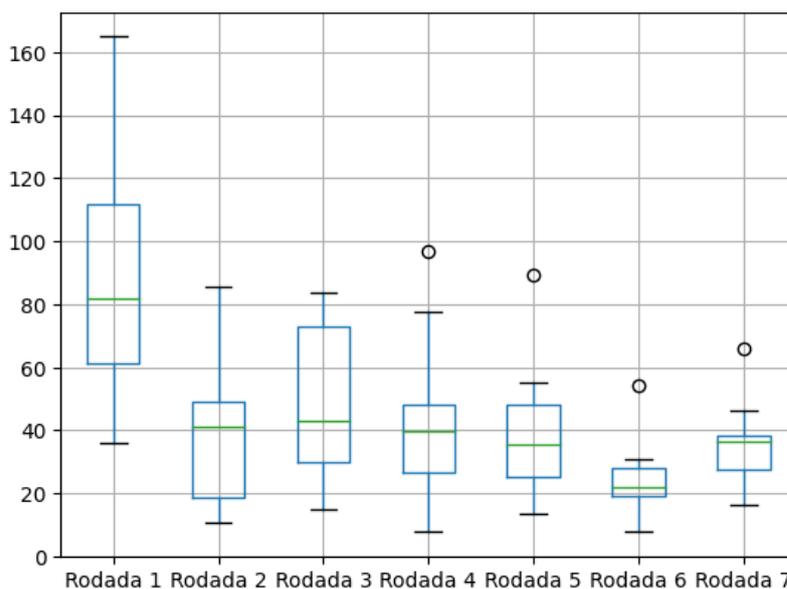
Tabela 7 - Estatísticas descritivas do grupo homens com conhecimento

	Rodada 1	Rodada 2	Rodada 3	Rodada 4	Rodada 5	Rodada 6	Rodada 7
count	13.000000	13.000000	13.000000	13.000000	13.000000	13.000000	13.000000
mean	93.859231	39.460000	48.100000	42.390769	38.786923	23.693846	34.668462
std	41.843574	22.62657	23.539771	25.288180	20.138162	11.449614	12.879802
min	36.180000	10.880000	15.200000	7.740000	13.480000	7.880000	16.220000
25%	61.450000	18.840000	29.770000	26.800000	25.370000	19.260000	27.690000
50%	82.010000	41.080000	43.170000	39.650000	35.340000	21.980000	36.300000
75%	111.560000	49.100000	73.100000	48.240000	48.100000	27.930000	38.280000
max	164.860000	85.660000	83.560000	96.660000	89.230000	54.030000	65.900000

Fonte: O autor (2024).

A primeira rodada desse grupo tem média de aproximadamente 93,85; com uma amplitude de 128,68 e desvio padrão de 41,84, o que sugere uma distribuição assimétrica com assimetria a direita, além de uma dispersão considerável dos dados. Outro aspecto importante são os quartis, onde temos que 25% dos dados estão abaixo de 61,45, 50% desses dados estão abaixo de 82,01 e 75% estão abaixo de 111,50, a diferença entre os quartis é considerável, já que entre o primeiro e o terceiro existe uma diferença de 50,05, afirmando a suposição de assimetria e grande dispersão dos dados. A Figura 14 traz a visualização dos dados através do diagrama boxplot.

Figura 14 - Diagrama Boxplot para o grupo homens com conhecimento



Fonte: O autor (2024).

Já na segunda rodada a média do tempo de execução é de 39,46, com amplitude de 70,6 e desvio padrão de 22,62, apesar da média, desvio padrão e amplitude dessa rodada serem menores que os dados da rodada 1, ainda existe uma alta variação. Analisando os quartis em conjunto com o boxplot da Figura 14, vemos que não existe a presença de outliers, sugerindo que o alto desvio padrão não está relacionado com valores pontuais que destoam significativamente, mas sim da própria variação dos dados.

A terceira rodada apresenta média de 48,1, amplitude de 68,36 e desvio padrão de 23,53, analisando a Figura 14 é possível ver que a média é maior que a mediana sugerindo que existe uma assimetria a direita. Outro aspecto importante que pode ser constatado é a dispersão dos dados, devido a amplitude e o desvio padrão relativamente alto, quando comparado com a média. Analisando os quartis, vemos que 25% dos dados são menores que 29,77, 50% dos dados são menores que 43,17 e 75% dos dados são menores que 73,1.

A quarta rodada apresenta média de 42,39, amplitude de 88,92 e desvio padrão de 25,28. Os quartis apresentam que 25% dos dados são menores que 26,80, 50% dos dados são menores que 39,65 e 75% dos dados são menores que 48,24, sugerindo que existe uma alta variabilidade nas distribuições dos dados, indicando a presença de outlier, que é confirmado pela Figura 14.

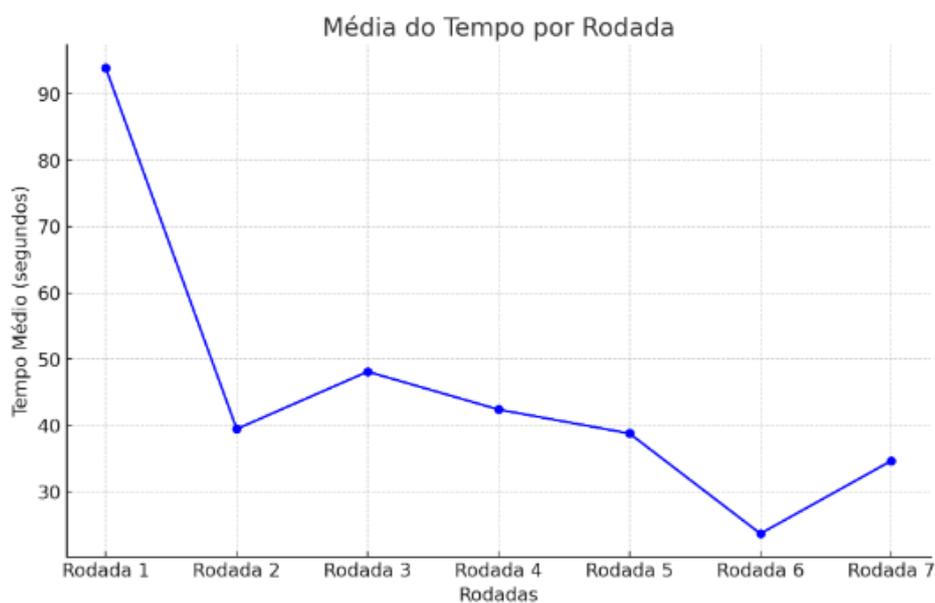
A quinta rodada possui média de 38,78, amplitude de 75,75 e desvio padrão de 20,13. Os quartis indicam que 25% dos dados são menores que 25,37, 50% são menores que 35,34 e 75% dos dados são menores que 48,10. Os elevados valores do desvio padrão e da amplitude, quando comparado a média, sugerem a existência de outlier o que é confirmado pela Figura 14.

A sexta rodada é a que apresenta menor dispersão dos dados, apesar de existir outlier, como pode ser visto na Figura 14. Essa rodada apresenta média de 23,69, desvio padrão de 11,44 e amplitude de 46,15. Os quartis indicam que 25% dos dados estão abaixo de 19,26, 50% dos dados estão abaixo de 21,98 e 75% dos dados são menores que 27,93.

A sétima rodada também tem uma baixa dispersão dos dados, apesar de também apresentar outliers, no entanto não tão baixa quanto a sexta rodada. A rodada sete apresenta média de 34,66, desvio padrão de 12,87 e amplitude de 49,68, já os quartis apresentam que 25% dos dados são menores que 16,22, 50% dos dados estão abaixo de 36,30 e 75% dos dados

Por fim, a Figura 15, traz o tempo médio de duração de cada rodada para o grupo homens com conhecimento.

Figura 15 - Tempo média de duração das rodadas para o grupo homens com conhecimento



Fonte: O autor (2024).

4.1.1.2. Estatística descritiva do grupo mulheres com conhecimento

As estatísticas descritivas para o grupo de mulheres com conhecimento podem ser vistas no Tabela 8, nesse está descrito a média, desvio padrão, valores máximos e mínimo e os quartis.

Tabela 8 - Estatísticas descritivas do grupo mulheres com conhecimento.

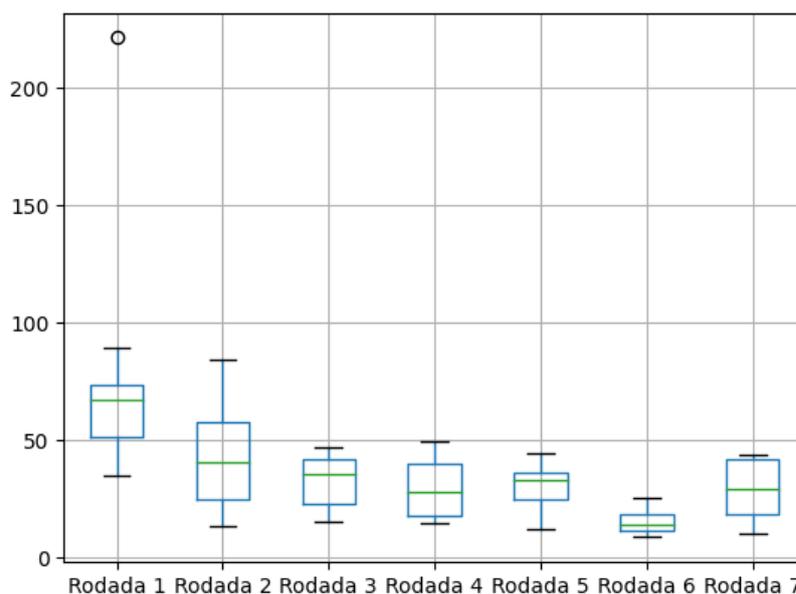
	Rodada 1	Rodada 2	Rodada 3	Rodada 4	Rodada 5	Rodada 6	Rodada 7
count	8.000000	8.000000	8.000000	8.000000	8.000000	8.000000	8.000000
mean	80.265000	43.273750	32.807500	29.580000	29.907500	15.050000	28.395000
std	59.691518	24.331122	11.595802	14.062762	10.763297	5.435579	13.882056
min	34.880000	12.940000	14.990000	14.500000	11.910000	8.480000	9.740000
25%	51.430000	24.565000	22.777500	17.807500	24.437500	11.230000	18.575000
50%	67.235000	40.555000	35.395000	27.730000	32.900000	13.990000	29.035000
75%	73.360000	57.675000	41.872500	39.575000	36.125000	18.402500	41.867500
max	221.510000	84.180000	47.060000	49.190000	44.440000	25.230000	43.460000

Fonte: O autor (2024).

A primeira rodada apresenta uma média de 80,27, com uma amplitude de 186,63 e desvio padrão de 56,73, resultando em um coeficiente de variação de 74,37%, o que evidencia uma alta dispersão relativa dos dados. A análise dos quartis mostra que 25% dos tempos estão

abaixo de 51,43, 50% estão abaixo de 67,30, e 75% abaixo de 73,36, sugerindo que a maioria dos tempos estão concentrados na faixa inferior, contudo a presença de outlier) estão influenciando a média e a amplitude. A Figura 16, com o boxplot dessa rodada, evidencia a presença de outliers significativos.

Figura 16 - Diagrama boxplot para o grupo mulheres com conhecimento.



Fonte: O autor (2024).

Na segunda rodada, a média é de 43,27, com amplitude de 71,24 e desvio padrão de 20,66, resultando em um coeficiente de variação de 56,23%, menor que o da rodada 1, mas ainda indicando considerável dispersão. Os quartis indicam que 25% dos tempos estão abaixo de 24,57, 50% abaixo de 46,22, e 75% abaixo de 57,68. O boxplot na Figura 16 mostra a ausência de outliers extremos, sugerindo que a alta variação é decorrente da distribuição natural dos dados.

A terceira rodada apresenta uma média de 32,81, amplitude de 36,73, e desvio padrão de 12,83, resultando em um coeficiente de variação de 35,34%, o menor até agora, sugerindo maior homogeneidade nos tempos. Os quartis mostram que 25% dos dados estão abaixo de 22,78, 50% abaixo de 32,13, e 75% abaixo de 41,87, confirmando uma dispersão mais controlada. A Figura 16 evidencia que não há outliers significativos.

Na quarta rodada, a média é de 29,58, com amplitude de 40,71 e desvio padrão de 14,74, resultando em um coeficiente de variação de 47,54%. Os quartis mostram que 25% dos tempos estão abaixo de 17,81, 50% abaixo de 26,80, e 75% abaixo de 39,58. Apesar da relativa

dispersão, não há valores que se destaquem como outliers significativos no boxplot da Figura 16.

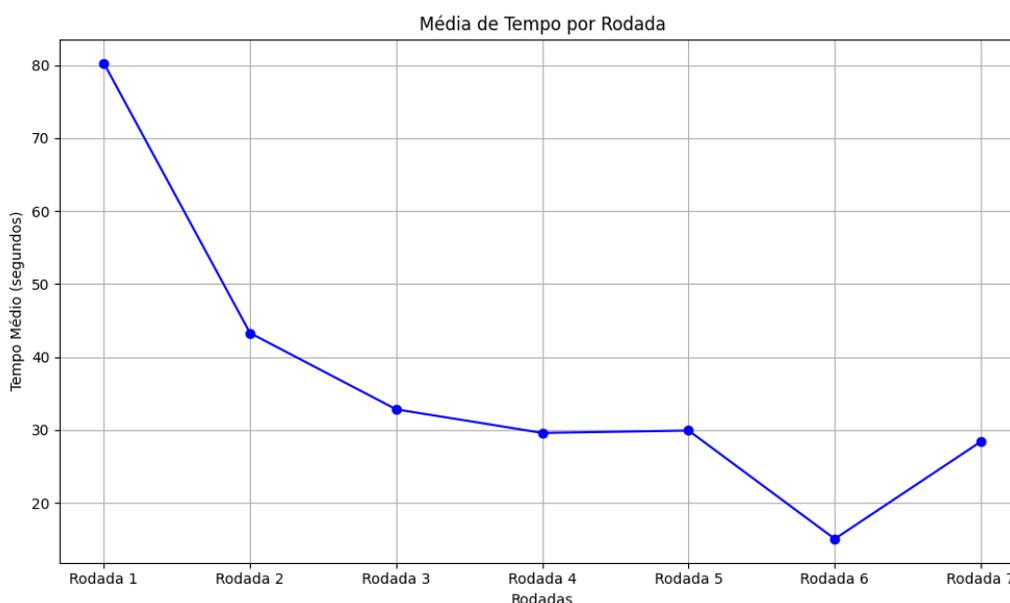
A quinta rodada possui uma média de 29,91, amplitude de 32,53, e desvio padrão de 10,90, com um coeficiente de variação de 35,99%, semelhante ao da rodada 3. Os quartis indicam que 25% dos tempos estão abaixo de 24,44, 50% abaixo de 31,64, e 75% abaixo de 36,13, sugerindo uma distribuição mais equilibrada. Não há outliers aparentes na Figura 16.

A sexta rodada apresenta a menor média entre todas, 15,05, com amplitude de 15,08 e desvio padrão de 5,98, resultando em um coeficiente de variação de 36,12%. Os quartis mostram que 25% dos dados estão abaixo de 11,23, 50% abaixo de 18,47, e 75% abaixo de 18,40, indicando baixa dispersão relativa e consistência nos tempos. Não há outliers visíveis na Figura 16.

Por fim, a sétima rodada tem uma média de 28,40, amplitude de 33,72, e desvio padrão de 8,75, com um coeficiente de variação de 48,89%, intermediário entre as rodadas anteriores. Os quartis indicam que 25% dos tempos estão abaixo de 18,58, 50% abaixo de 25,41, e 75% abaixo de 41,87. A Figura 16 confirma a ausência de outliers e uma dispersão relativamente moderada.

A Figura 17 expressa o gráfico com o tempo médio da duração das rodadas, para o grupo de mulheres com conhecimento.

Figura 17 - Tempo médio de duração das rodadas para o grupo mulheres com conhecimento



Fonte: O autor (2024).

4.1.1.3. Estatística descritiva do grupo homens sem conhecimento

As estatísticas descritivas para o grupo de homens sem conhecimento podem ser vistas no Tabela 9, nesse está descrito a média, desvio padrão, valores máximos e mínimo e os quartis.

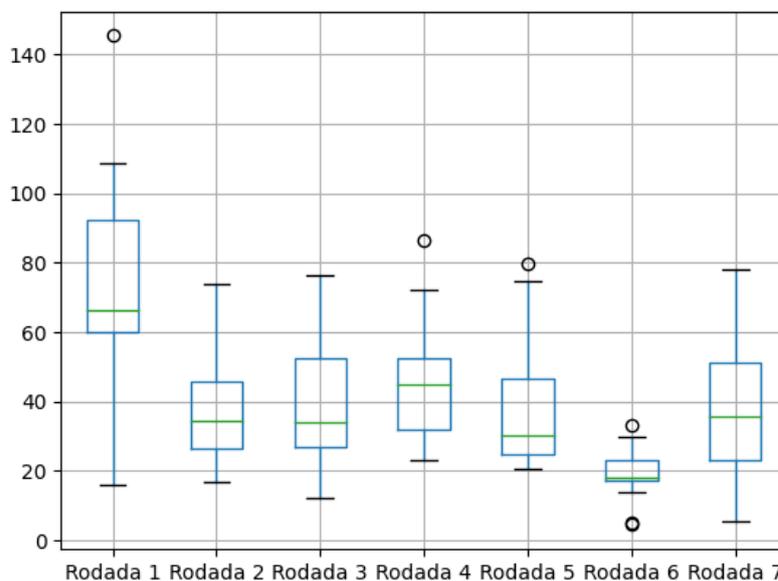
Tabela 9 - Estatísticas descritivas do grupo homens sem conhecimento.

	Rodada 1	Rodada 2	Rodada 3	Rodada 4	Rodada 5	Rodada 6	Rodada 7
count	15.000000	15.000000	15.000000	15.000000	15.000000	15.000000	15.000000
mean	74.573333	38.105333	39.712667	45.356000	38.430000	19.374667	37.642000
std	31.578198	16.618089	18.398557	18.266879	18.802112	7.929935	19.50773
min	15.900000	16.960000	12.210000	22.910000	20.760000	4.420000	5.410000
25%	59.905000	26.590000	26.820000	31.675000	24.770000	17.380000	23.080000
50%	66.370000	34.250000	34.070000	44.720000	30.290000	18.160000	35.730000
75%	92.185000	45.555000	52.310000	52.505000	46.350000	23.205000	51.235000
max	145.310000	73.700000	76.290000	86.450000	79.610000	33.160000	77.980000

Fonte: O autor (2024).

A primeira rodada para o grupo de homens sem conhecimento, tem média de aproximadamente 74,57, com uma amplitude de 129,41 e desvio padrão de 29,69, o que sugere uma distribuição assimétrica com assimetria à direita, além de uma dispersão considerável dos dados. Outro aspecto importante são os quartis: 25% dos dados estão abaixo de 59,91, 50% desses dados estão abaixo de 66,37, e 75% estão abaixo de 92,19. A diferença entre o primeiro e o terceiro quartil é de 32,28, indicando uma dispersão significativa dos dados. A Figura 18 traz a visualização dessa rodada por meio de um boxplot, o gráfico apresenta a existência de outliers, reforçando a assimetria.

Figura 18 - Diagrama Boxplot para o grupo homens sem conhecimento



Fonte: O autor (2024).

Na segunda rodada, a média do tempo de execução é de 36,06, com uma amplitude de 80,61 e desvio padrão de 15,19. Apesar de a média, o desvio padrão e a amplitude dessa rodada serem menores que os da rodada 1, ainda há uma variação notável. Analisando os quartis em conjunto com o boxplot da Figura 18, vemos que 25% dos dados estão abaixo de 25,12, 50% abaixo de 34,25, e 75% abaixo de 45,67, não havendo evidência de outliers nessa rodada. A variação parece ser consistente com a dispersão natural dos dados.

A terceira rodada apresenta média de 39,06, amplitude de 70,88 e desvio padrão de 15,47. Na Figura 18, é possível observar que a média é maior que a mediana (34,07), sugerindo uma leve assimetria à direita. Os quartis indicam que 25% dos dados são menores que 26,77, 50% menores que 34,07 e 75% menores que 59,63, o que reforça a ideia de dispersão moderada.

A quarta rodada apresenta média de 41,92, amplitude de 82,03 e desvio padrão de 16,67. Os quartis mostram que 25% dos dados são menores que 28,99, 50% são menores que 40,44 e 75% dos dados estão abaixo de 55,69. Embora os valores indiquem uma variação significativa, o boxplot da Figura 18 revela a presença de outliers, que contribuem para a maior amplitude dos dados.

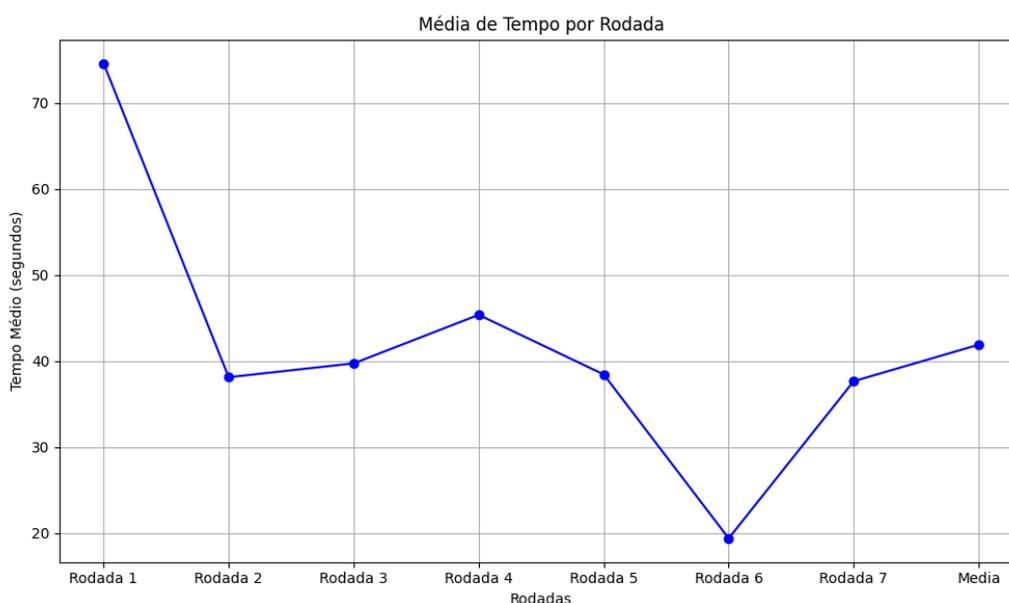
A quinta rodada possui média de 41,44, amplitude de 58,85 e desvio padrão de 19,79. Os quartis indicam que 25% dos dados são menores que 23,10, 50% menores que 45,14, e 75% menores que 47,33. A elevada amplitude e o desvio padrão sugerem a possível influência de outliers, como indicado no boxplot da Figura 18.

A sexta rodada apresenta a menor dispersão dos dados, com uma média de 21,4, desvio padrão de 7,64 e amplitude de 24,58. Os quartis indicam que 25% dos dados estão abaixo de 17,18, 50% abaixo de 21,68, e 75% abaixo de 24,59. Apesar de os valores indicarem uma alta consistência, o boxplot da Figura 18 confirma a presença de outliers.

Por fim, a sétima rodada também apresenta baixa dispersão dos dados, embora maior que a sexta rodada. A média é de 36,57, o desvio padrão é de 18,36, e a amplitude é de 72,57. Os quartis mostram que 25% dos dados estão abaixo de 21,58, 50% abaixo de 30,80, e 75% abaixo de 52,11. Apesar da relativa consistência, há indícios de outliers, confirmados pelo boxplot da Figura 18

A Figura 19 expressa o gráfico com o tempo médio da duração das rodadas, para o grupo de homens sem conhecimento. A comparação entre as rodadas evidencia que a primeira rodada apresenta a maior dispersão relativa e absoluta, influenciada por outliers extremos. Já as rodadas sexta e sétima mostram os tempos mais consistentes, com baixa amplitude e coeficientes de variação moderados. O uso do coeficiente de variação trouxe insights adicionais sobre a variabilidade relativa, complementando a análise com desvio padrão. Essa métrica é especialmente útil para comparar a homogeneidade entre rodadas com médias muito diferentes.

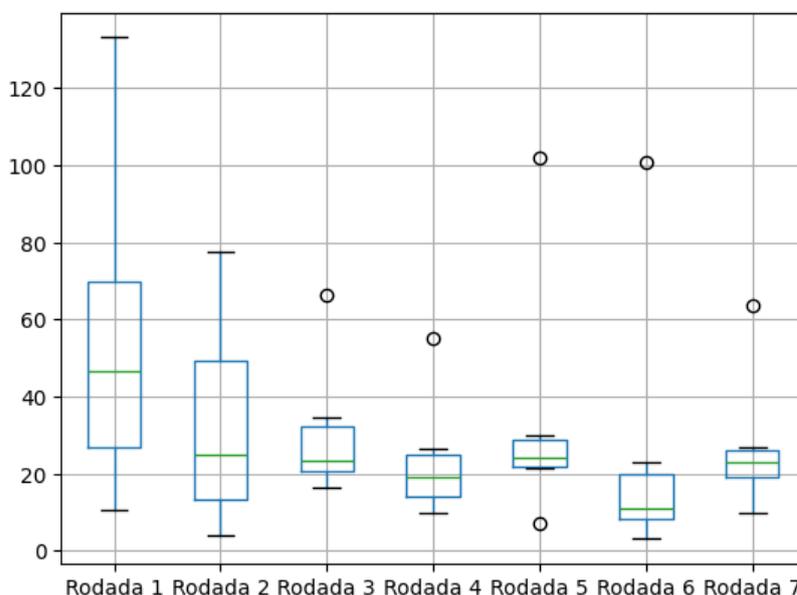
Figura 19 - Tempo média de duração das rodadas para o grupo homens sem conhecimento



4.1.1.4. Estatística descritiva do grupo mulheres sem conhecimento

A primeira rodada apresenta uma média de 55,78, com amplitude de 122,50 (diferença entre o máximo de 133,12 e o mínimo de 10,62) e desvio padrão de 44,25, resultando em uma alta variabilidade relativa. O coeficiente de variação, calculado como a razão entre o desvio padrão e a média, é 79,34%, reforçando a grande dispersão dos tempos. Os quartis mostram que 25% dos tempos estão abaixo de 27,01, 50% abaixo de 46,50 (mediana), e 75% abaixo de 69,71, enquanto o boxplot da Figura 20 revela a presença de outliers que impactam a distribuição dos dados.

Figura 20 - Diagrama Boxplot para o grupo mulheres sem conhecimento



Fonte: O autor (2024).

Na segunda rodada, a média é de 33,02, com amplitude de 73,68 e desvio padrão de 28,40, resultando em um coeficiente de variação de 85,99%, indicando uma alta variabilidade relativa. Os quartis indicam que 25% dos tempos estão abaixo de 13,23, 50% abaixo de 24,75 (mediana), e 75% abaixo de 49,33. O boxplot na Figura 20 sugere a presença de valores extremos, mas a maioria dos tempos está concentrada em torno da mediana.

A terceira rodada apresenta uma média de 30,70, com amplitude de 49,73 e desvio padrão de 18,44, resultando em um coeficiente de variação de 60,06%. Os quartis indicam que 25% dos tempos estão abaixo de 20,74, 50% abaixo de 23,33, e 75% abaixo de 32,21. A análise do boxplot mostra uma distribuição mais controlada, apesar da existência de outliers.

A quarta rodada tem uma média de 23,73, com amplitude de 45,01 e desvio padrão de 16,41, resultando em um coeficiente de variação de 69,13%. Os quartis mostram que 25% dos

dados estão abaixo de 13,99, 50% abaixo de 19,23 (mediana), e 75% abaixo de 24,81. Apesar da dispersão considerável, não há outliers evidentes na distribuição apresentada pelo boxplot.

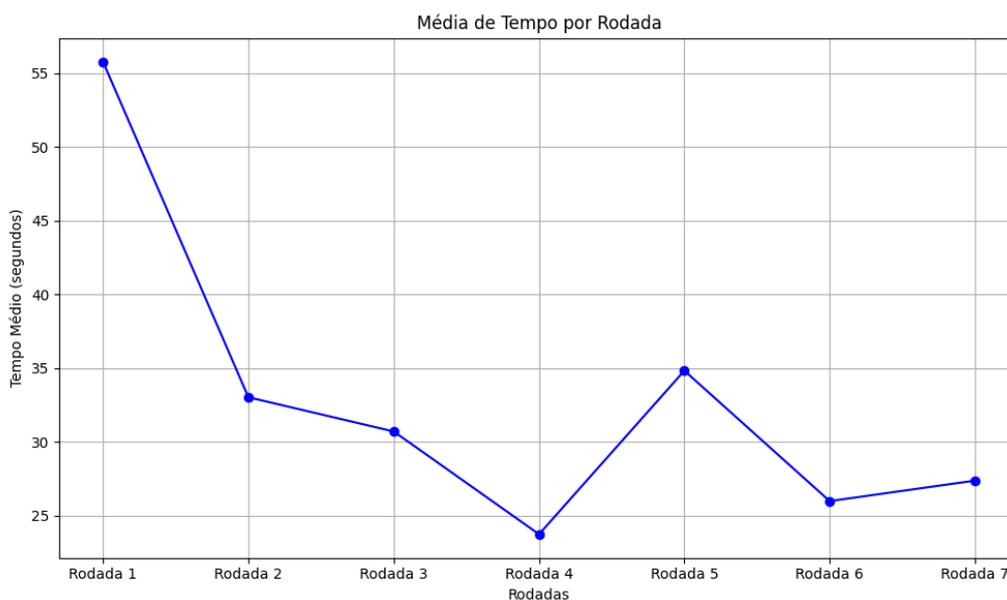
A quinta rodada possui a maior média entre as rodadas, 34,83, com amplitude de 94,66 e desvio padrão de 33,75, indicando uma alta variabilidade. O coeficiente de variação é de 96,96%, o maior dentre todas as rodadas. Os quartis indicam que 25% dos tempos estão abaixo de 21,69, 50% abaixo de 24,23, e 75% abaixo de 28,96, enquanto o boxplot revela a presença de valores extremos que contribuem para a dispersão.

A sexta rodada apresenta uma média de 25,98, com amplitude de 97,64 e desvio padrão de 37,20, resultando em um coeficiente de variação de 143,26%, o mais alto entre todas as rodadas, indicando alta dispersão relativa. Os quartis mostram que 25% dos tempos estão abaixo de 8,17, 50% abaixo de 10,82 (mediana), e 75% abaixo de 20,08. O boxplot da Figura 12 confirma a presença de outliers extremos, que impactam a distribuição dos tempos.

Por fim, a sétima rodada apresenta uma média de 27,37, amplitude de 53,69, e desvio padrão de 18,70, resultando em um coeficiente de variação de 68,32%. Os quartis mostram que 25% dos tempos estão abaixo de 18,94, 50% abaixo de 22,97, e 75% abaixo de 26,14. O boxplot evidencia uma distribuição relativamente homogênea, sem presença de outliers marcantes.

A Figura 21 apresenta o gráfico com o tempo médio da duração das rodadas, para o grupo de mulheres sem conhecimento. A análise dos dados revela que as rodadas quinta e sexta apresentam as maiores dispersões relativas, com coeficientes de variação superiores a 95%, influenciados por outliers significativos e alta amplitude. Em contrapartida, a terceira rodada possui uma distribuição mais controlada, com um coeficiente de variação menor e ausência de valores extremos. A comparação entre os quartis e as medidas de dispersão destaca a importância de observar tanto os valores centrais quanto os extremos para identificar tendências e possíveis anomalias nos dados.

Figura 21 - Tempo média de duração das rodadas para o grupo mulheres sem conhecimento



Fonte: O autor (2024).

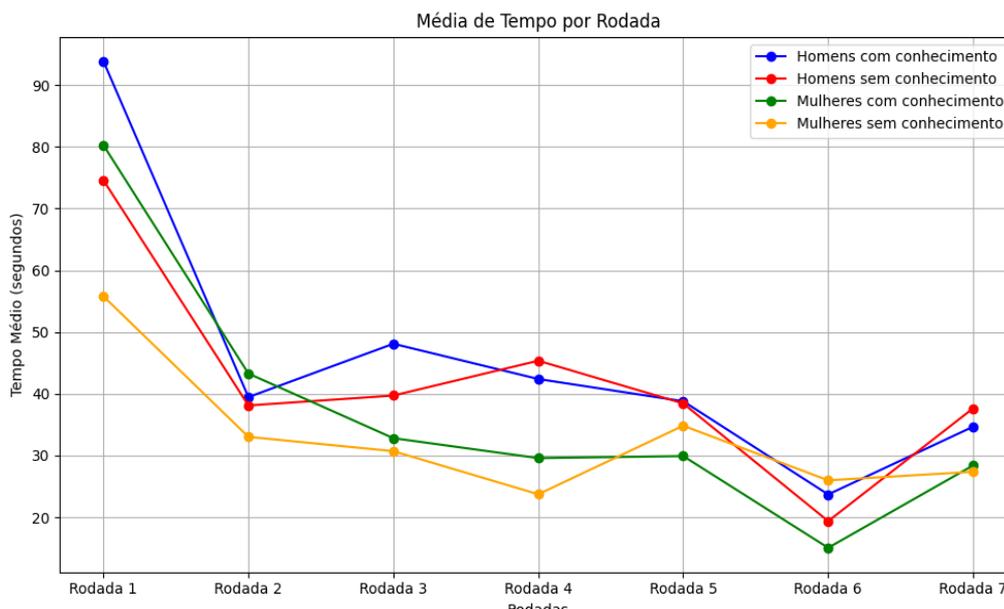
4.1.1.5. Comparação entre os grupos

Em termos de média, o grupo de Homens com Conhecimento apresenta os maiores valores na primeira rodada, enquanto o grupo de Mulheres com Conhecimento também têm uma média alta na primeira rodada, mas com uma grande variação nas rodadas seguintes. Os grupos de Homens sem Conhecimento e Mulheres sem Conhecimento apresentam médias mais baixas, com destaque para a sexta Rodada, onde os valores de todos os grupos caem significativamente.

Quanto ao desvio padrão quanto coeficiente de variação, os grupos de Mulheres com Conhecimento e de Mulheres sem Conhecimento apresentam os maiores valores nas rodadas iniciais, o que sugere uma maior dispersão nos dados e a presença de outliers. O grupo de Homens com Conhecimento também apresenta grandes variações, mas principalmente na primeira Rodada. Já os Homens sem Conhecimento têm uma dispersão mais controlada, com os dados mais concentrados em torno da média, especialmente nas rodadas finais.

Todos os grupos apresentam assimetria, mas os grupos de mulheres (com e sem conhecimento) parecem ter uma dispersão mais acentuada, com outliers impactando as médias e desvios padrão, dessa maneira a primeira rodada e a sétima rodada são particularmente influenciadas por valores extremos em todos os grupos. A Figura 22, traz a comparação das médias do tempo de execução de cada rodada e para todos os grupos

Figura 22 - Comparação entre as médias das rodadas dos grupos



Fonte: O autor (2024).

De maneira geral, o grupo de Homens com Conhecimento tende a apresentar uma maior variabilidade, especialmente na primeira rodada. Já os grupos de Mulheres (com e sem conhecimento) mostram uma grande variação nos dados, com outliers que afetam significativamente as médias e os desvios padrão. O grupo de Homens sem Conhecimento tem menor dispersão, com os dados mais concentrados em torno da média, especialmente nas rodadas finais. Cada grupo exibe características distintas, com variações significativas dependendo do nível de conhecimento e da rodada analisada, refletindo a diversidade no desempenho de cada grupo nas diferentes fases da avaliação intracritério.

4.1.2. Testes paramétricos e não paramétricos

A Análise de variância (ANOVA), é um teste estatístico utilizado para avaliar a existência de uma diferença, estatisticamente significativa, entre as médias de grupos independentes. O teste foi desenvolvido no final do século 19 e início do século 20, pelo estatístico Ronald Fisher. Segundo Fisher (1925), a ANOVA se baseia na comparação entre a variabilidade dos dados dos grupos com variabilidade dos dados entre cada grupo. O objetivo geral do teste é verificar se a diferença nas médias é decorrente de variações aleatórias ou efeitos sistemáticos. Howell (2012) explica que, a ANOVA é baseado em um teste de hipótese que testa a hipótese nula, H_0 , que defende a igualdade entre as médias, contra a hipótese alternativa,

H_1 , que contesta essa igualdade afirmando que as médias são diferentes. Sokal e Rohlf (1995) abordam a importância da ANOVA no contexto experimental, já que ela permite a decomposição da variabilidade total, Kutner *et al.* (2005) contribuem afirmando que, ao teste é um modelo eficaz para avaliar, a significância estatística, dos efeitos de fatores categóricos tem sobre uma variável resposta.

A análise de variância é ideal para tratar os casos que há mais de dois grupos envolvidos no experimento e consequentemente evitando os erros do tipo 1 e 2, que são descritos no Quadro 1. Apesar da ANOVA se mostrar robusta, flexível e eficaz, Ferguson (2009) discute a existência de pressupostos para a aplicação da ANOVA, assim como sua importância para a validade dos testes de hipótese, os pressupostos para aplicação são descritos no Quadro 2.

Quadro 1 - Erros do Tipo 1 e Tipo 11

ERRO	DESCRIÇÃO	CONSEQUENCIA
Tipo 1, falso positivo.	Rejeitar a hipótese nula verdadeira.	Concluir que existe um efeito, quando não há.
Tipo 2, Falso negativo.	Não rejeitar a hipótese nula falsa.	Não detectar um efeito real.

Fonte: O autor (2024).

Quadro 2 - Pressupostos da ANOVA

PRESSUPOSTO	DESCRIÇÃO	TESTE
Independência dos dados	As observações não devem se influenciar	Não existe teste formal, deve ser
Normalidade	Os resíduos devem ser normalmente distribuídos	Teste de Shapiro-Wilk
Homoceadasticidade	Variância constante dos resíduos	Teste de Breusch-Pagan

Fonte: O autor (2024).

Apesar da existência dos pressupostos na aplicação do teste ANOVA, necessários para garantir a qualidade dos resultados, existem alternativas que podem ser utilizadas como sugerido por Montgomery (2017). Em situações em que um ou mais pressupostos não são atendidos, alternativas conhecidas como testes não paramétricos, podem ser utilizados como formas de substituir o teste ANOVA sem perder a qualidade nas análises. A Tabela 10 apresenta os tipos de ANOVA e os testes não paramétricos correspondentes, contudo as análises de variância que operam com mais de uma via em estado de repetição um pressuposto a mais

precisa ser atendido, a esfericidade, que está relacionada com a homogeneidade das variâncias das diferenças e é testado a partir do teste de Mauchly.

Tabela 10 - Tipos de ANOVA e seus correspondentes não paramétricos.

Tipo de ANOVA	Proposito	Teste não paramétrico correspondente
ANOVA de uma via.	Compara as médias entre três ou mais grupos independentes.	Teste de Kruskal-Wallis
ANOVA de duas vias	Compara as médias de duas variáveis independentes.	Teste de Kruskal-Wallis para cada fator
ANOVA com medidas repetidas	Compara médias entre três ou mais condições repetidas.	Teste de Friedman
ANOVA de duas vias com medidas repetidas	Compara médias entre duas variáveis independentes com medições repetidas.	Teste de Friedman estendido
ANOVA mista	Combina variáveis dependentes e independentes em um único modelo	Teste Permutacional
ANCOVA	Compara as médias partindo de uma variável contínua.	ANCOVA não paramétrica
MANOVA	Compara as médias para múltiplas variáveis dependentes	MANOVA não paramétrica

Fonte: O autor (2024).

Como já apresentado anteriormente, para garantir a qualidade dos resultados é preciso que os pressupostos sejam atendidos, além disso é necessário escolher qual dos testes ANOVA, apresentados na Tabela 10, melhor se ajusta ao escopo desse trabalho. Devido a natureza do experimento, onde o participante é submetido a avaliação de sete critérios de forma individual e subsequente, sem a influência de outros participantes, avalia-se que o teste que mais se ajusta aos dados é a ANOVA com medida repetidas. Esse teste sugere que um único fator é avaliado diversas vezes para cada observação, nesse caso, o fator é o tempo de execução da rodada e as medidas são as rodadas que compõem o experimento.

Para analisar os dados referentes ao grupo de homens com conhecimento, os dados de todas as rodadas foram submetidos aos testes de Shapiro-Wilk, que testa a normalidade dos dados a partir do teste de hipótese onde:

- H_0 : Os dados seguem uma distribuição normal.
- H_1 : Os dados não seguem uma distribuição normal.

Para avaliar é preciso comparar o p-valor encontrado com o nível de significância, que nesse caso é de 0,05, se o valor de p for maior que o nível de significância os dados seguem uma distribuição normal, ou seja, aceitamos a hipótese nula. Por outro lado, se o valor de p for menor que o nível de significância os dados não seguem uma distribuição normal e rejeitamos a hipótese nula. De forma análoga, o teste de Breunsch-Pagan para a homocedasticidade é realizado. A Tabela 11 **Error! Reference source not found.** apresenta o p-valor para os testes de normalidade e homoscedasticidade, assim como a interpretação dos testes.

Tabela 11 - Teste de normalidade e homocedásticos para o grupo Homem com conhecimento

Rodada	Shapiro-Wilk	Interpretação	Breusch-Pagan	Interpretação
1°	p-valor = 0,362	Normalidade	p-valor = 0.975	Homoscedasticidade
2°	p-valor = 0.421	Normalidade	p-valor = 0.066	Homoscedasticidade
3°	p-valor = 0.244	Normalidade	p-valor = 0.124	Homoscedasticidade
4°	p-valor = 0.571	Normalidade	p-valor = 0.119	Homoscedasticidade
5°	p-valor = 0.192	Normalidade	p-valor = 0.087	Homoscedasticidade
6°	p-valor = 0.060	Normalidade	p-valor = 0.184	Homoscedasticidade
7°	p-valor = 0.242	Normalidade	p-valor = 0.114	Homoscedasticidade

Fonte: O autor (2024).

Analisando os resultados obtidos na Tabela 11, é possível concluir que os resultados de todas as rodadas seguem os pressupostos de normalidade e homoscedasticidade, dessa maneira, possibilitando a aplicação da ANOVA de uma via com repetição para o grupo homens com conhecimento.

Tabela 12 - Teste de normalidade e Homoscedasticidade para o grupo mulheres com conhecimento

Rodada	Shapiro-Wilk	Interpretação	Breusch-Pagan	Interpretação
1°	p-valor = 0,051	Normalidade	p-valor = 0.439	Homoscedasticidade
2°	p-valor = 0.844	Normalidade	p-valor = 0.420	Homoscedasticidade
3°	p-valor = 0.494	Normalidade	p-valor = 0.597	Homoscedasticidade
4°	p-valor = 0.652	Normalidade	p-valor = 0.296	Homoscedasticidade
5°	p-valor = 0.192	Normalidade	p-valor = 0.238	Homoscedasticidade
6°	p-valor = 0.625	Normalidade	p-valor = 0.282	Homoscedasticidade
7°	p-valor = 0.200	Normalidade	p-valor = 0.256	Homoscedasticidade

Fonte: O autor (2024).

A Tabela 12 apresenta o resultado dos testes de normalidade e homoscedasticidade para o grupo mulheres com conhecimento. Analisando o p-valor das rodadas em cada um dos testes, é possível concluir que esse grupo também atende aos pressupostos.

Tabela 13 - Teste de normalidade e homoscedasticidade para o grupo Homem sem conhecimento

Rodada	Shapiro-Wilk	Interpretação	Breusch-Pagan	Interpretação
1°	p-valor = 0.917	Normalidade	p-valor = 0.209	Homoscedasticidade
2°	p-valor = 0.393	Normalidade	p-valor = 0.348	Homoscedasticidade
3°	p-valor = 0.273	Normalidade	p-valor = 0.177	Homoscedasticidade
4°	p-valor = 0.288	Normalidade	p-valor = 0.129	Homoscedasticidade
5°	p-valor = 0.121	Normalidade	p-valor = 0.135	Homoscedasticidade
6°	p-valor = 0.445	Normalidade	p-valor = 0.418	Homoscedasticidade
7°	p-valor = 0.987	Normalidade	p-valor = 0.171	Homoscedasticidade

Fonte: O autor (2024).

Os resultados dos testes para o grupo de homens sem conhecimento, podem ser vistos na Tabela 13 com base nas análises é possível concluir que esse grupo também atende aos pressupostos de normalidade e homoscedasticidade.

Tabela 14 - Teste de normalidade e homoscedasticidade para o grupo mulher sem conhecimento

Rodada	Shapiro-Wilk	Interpretação	Breusch-Pagan	Interpretação
1°	p-valor = 0.465	Normalidade	p-valor = 0.423	Homoscedasticidade
2°	p-valor = 0.488	Normalidade	p-valor = 0.423	Homoscedasticidade
3°	p-valor = 0.033	Não Normal	p-valor = 0.423	Homoscedasticidade
4°	p-valor = 0.069	Normalidade	p-valor = 0.423	Homoscedasticidade
5°	p-valor = 0.007	Não Normal	p-valor = 0.423	Homoscedasticidade
6°	p-valor = 0.019	Não Normal	p-valor = 0.423	Homoscedasticidade
7°	p-valor = 0.498	Normalidade	p-valor = 0.423	Homoscedasticidade

Fonte: O autor (2024).

O grupo mulheres sem conhecimento teve um comportamento diferente dos outros grupos, observando a Tabela 14 constatamos que três das sete rodadas apresentam não normalidade. Tabachnick e Fidell (2013) sugerem a utilização de uma transformação logarítmica para contornar a não normalidade das observações, a transformação diminui a

influência dos outliers estabilizando a variância dos dados, assim aproximando a distribuição dos dados da distribuição normal.

Aplicando a transformação logarítmica para a terceira rodada, quinta rodada e sexta rodada, e refazendo o teste de Shapiro-Wilk encontramos os novos valores para p de 0,261, 0,337 e 0,174, para as rodadas respectivamente. Os novos valores de p indicam que essas rodadas respeitam o pressuposto da normalidade dos dados.

4.1.2.1. Grupo homens com conhecimento

Agora que os pressupostos de normalidade, homoscedasticidade e independência das observações foram satisfeitos, é possível investigar a relação entre as rodadas. Antes de avançar nos discursões, mais um pressuposto precisa ser verificado, como o modelo de ANOVA com medidas repetidas foi o mais adequado para o experimento, ele exige que a esfericidade seja entendida. A esfericidade é a condição da diferença entre as variâncias de todos os pares possíveis, dentre as variáveis independentes, sejam iguais, contudo, a violação desse pressuposto ocorre uma distinção nos cálculos das variâncias. Mouchly (1940), apresentou um teste que verifica a condição de esfericidade, posteriormente esse teste foi batizado como Teste de esfericidade de Mouchly ou W de Mouchly, Figura 23 apresenta o teste de esfericidade de Mouchly para o grupo de homens com conhecimento.

Figura 23 - Teste de esfericidade para o grupo homens com conhecimento.

Teste de esfericidade de Mauchly^a

Medir: MEASURE_1

Efeito entre assuntos	W de Mauchly	Aprox. Qui-quadrado	df	Sig.	Epsilon ^b		
					Greenhouse-Geisser	Huynh-Feldt	Limite inferior
rodadas	,008	46,710	20	,001	,534	,752	,167

Fonte: O autor (2024).

Observando o resultado do teste, vemos que o valor de 0,008 é menor que o nível de significância de 0,05, o que sugere a violação do da esfericidade. Em situações como essa é preciso corrigir a esfericidade, neste trabalho iremos usar o método desenvolvido por Greenhouse e Geisser (1959), esse método atua na redução dos graus de liberdade induzindo uma razão F que minimize o erro do tipo I, contudo a relação F real não muda aplicando a correção. A Figura 24 mostra o resultado da ANOVA com repetição.

Figura 24 - Resultado da ANOVA corrigido para grupo homens com conhecimento.

Testes de efeitos entre assuntos

Medir: MEASURE_1

Origem		Tipo III Soma dos Quadrados	df	Quadrado Médio	Z	Sig.
rodadas	Esféricidade considerada	39371,230	6	6561,872	15,105	,000
	Greenhouse-Geisser	39371,230	3,205	12283,060	15,105	,000

Fonte: O autor (2024).

Avaliando os resultados apresentados na Figura 24, vemos que a o p-valor foi menor que 0,001, indicando que existe uma diferença estatisticamente significativa entre as rodadas. Outro dado que corrobora para a interpretação de diferença significativa e o valor de Z, que para o grupo homens com conhecimento obteve um valor de 15,105 o que sugere uma diferença substancial.

A Tabela 15 apresenta o resultado dos testes de contraste, que avaliam as formas que as variações dos tempos de execução entre as rodadas podem ser descritos, tendo como intuito entender a estrutura da variação. Na Tabela 15 são testadas seis formas de descrever a variação do tempo ao longo das rodadas, para o grupo homens com conhecimento, a variação pode ser descrita como linear, quadrática, cúbica e de ordem 4, devido ao p-valor de cada uma ser menor que 0,05. Tendo isso em vista, concluímos que existe um padrão nas variações e esses padrões podem ser descritos nas formas que foram estatisticamente significativas.

Tabela 15 - Teste de contraste para o grupo homens com conhecimento

Testes de contrastes entre assuntos

Medir: MEASURE_1

Origem		Tipo III Soma dos Quadrados	df	Quadrado Médio	Z	Sig.
rodadas	Linear	22149,348	1	22149,348	31,196	,000
	Quadrático	6982,852	1	6982,852	16,986	,001
	Cúbico	2521,127	1	2521,127	23,645	,000
	Ordem 4	6844,040	1	6844,040	10,397	,007
	Ordem 5	282,064	1	282,064	,680	,426
	Ordem 6	591,798	1	591,798	1,934	,190
Erro(rodadas)	Linear	8520,161	12	710,013		
	Quadrático	4933,185	12	411,099		
	Cúbico	1279,502	12	106,625		
	Ordem 4	7899,431	12	658,286		

Ordem 5	4974,333	12	414,528	
Ordem 6	3671,947	12	305,996	

Fonte: O autor (2024).

Anteriormente foi comprovado a existência de uma diferença significativa entre as médias, porém uma nova questão surge “*Se existe uma diferença entre as médias das rodadas, quais rodadas diferem entre si?*”, para responder a essa pergunta é preciso realizar um teste de post-hoc para comparar as médias de forma pareada. Para realizar a comparação entre os pares, utilizamos o teste sugerido por Sidak (1967), no teste de Sidak a comparação entre os pares de rodadas possível e se a comparação tiver um p-valor menor que 0,05 existe uma diferença estatisticamente significativa entre as médias do par analisado. Analisando o teste pós-hoc para as rodadas do grupo homens com conhecimento, temos que apenas a primeira rodada tem uma diferença estatística na média, quando comparada com as outras seis rodadas. Comparando o gráfico das médias para esse grupo, expressos na Figura 15, com o resultado obtido na Tabela 16, concluímos que apesar da segunda e sexta rodada sejam, visualmente distantes das outras rodadas, o teste comprova que não existe diferença significativa dessas duas rodadas com as outras, com a exceção da primeira.

Tabela 16 - Teste pós-hoc para o grupo homens com conhecimento.

Comparações de pares

Medir: MEASURE_1

(I) rodadas	(J) rodadas	Diferença média (I- J)	Erro Padrão	Sig. ^b	95% Intervalo de Confiança para Diferença ^b	
					Limite inferior	Limite superior
1	2	54,399 [*]	11,053	,007	12,175	96,623
	3	45,759 [*]	8,473	,003	13,392	78,126
	4	51,468 [*]	10,285	,006	12,181	90,756
	5	55,072 [*]	11,592	,010	10,790	99,354
	6	70,165 [*]	11,080	,001	27,842	112,489
	7	59,191 [*]	10,193	,002	20,254	98,128
2	1	-54,399 [*]	11,053	,007	-96,623	-12,175
	3	-8,640	7,686	,999	-38,001	20,721
	4	-2,931	10,333	1,000	-42,402	36,541
	5	,673	5,604	1,000	-20,733	22,079
	6	15,766	6,358	,461	-8,523	40,055
	7	4,792	7,763	1,000	-24,862	34,445
3	1	-45,759 [*]	8,473	,003	-78,126	-13,392

	2	8,640	7,686	,999	-20,721	38,001
	4	5,709	5,354	1,000	-14,745	26,163
	5	9,313	6,715	,988	-16,336	34,963
	6	24,406	6,995	,090	-2,315	51,127
	7	13,432	5,957	,608	-9,326	36,189
4	1	-51,468*	10,285	,006	-90,756	-12,181
	2	2,931	10,333	1,000	-36,541	42,402
	3	-5,709	5,354	1,000	-26,163	14,745
	5	3,604	9,325	1,000	-32,018	39,225
	6	18,697	7,381	,428	-9,500	46,894
	7	7,722	6,207	,997	-15,989	31,434
5	1	-55,072*	11,592	,010	-99,354	-10,790
	2	-,673	5,604	1,000	-22,079	20,733
	3	-9,313	6,715	,988	-34,963	16,336
	4	-3,604	9,325	1,000	-39,225	32,018
	6	15,093	7,389	,749	-13,134	43,320
	7	4,118	5,978	1,000	-18,716	26,953
6	1	-70,165*	11,080	,001	-112,489	-27,842
	2	-15,766	6,358	,461	-40,055	8,523
	3	-24,406	6,995	,090	-51,127	2,315
	4	-18,697	7,381	,428	-46,894	9,500
	5	-15,093	7,389	,749	-43,320	13,134
	7	-10,975	4,041	,328	-26,410	4,461
7	1	-59,191*	10,193	,002	-98,128	-20,254
	2	-4,792	7,763	1,000	-34,445	24,862
	3	-13,432	5,957	,608	-36,189	9,326
	4	-7,722	6,207	,997	-31,434	15,989
	5	-4,118	5,978	1,000	-26,953	18,716
	6	10,975	4,041	,328	-4,461	26,410

Fonte: O autor (2024).

4.1.2.2. Grupo mulheres com conhecimento

De forma análoga ao grupo homens com conhecimento, é preciso verificar o pressuposto da esfericidade através do teste de Mauchly. A Figura 25 apresenta o resultado do teste, dessa maneira, concluímos que o grupo mulheres com conhecimento não cumpri o pressuposto da esfericidade, uma vez que o p-valor é menor que o nível de significância de 0,05. Neste caso também será necessário aplicar a correção de Greenhouse-Geisser para obter um resultado significativo.

Figura 25 - Teste de esfericidade para o grupo mulheres com conhecimento.

Teste de esfericidade de Mauchly^a

Medir: MEASURE_1

Efeito entre assuntos	W de Mauchly	Aprox. Qui-quadrado	df	Sig.	Epsilon ^b		
					Greenhouse-Geisser	Huynh-Feldt	Limite inferior
Rodadas	,000	49,286	20	,001	,224	,258	,167

Fonte: O autor (2024).

A Figura 26 apresenta o resultado da ANOVA para o grupo de mulheres com conhecimento. Ao analisar a imagem vemos que o p-valor ajustado pelo método de Greenhouse-Geisser é de 0,041 que é menor que o nível de significância de 0,05, sugerindo que existe uma significância estatística entre as médias das rodadas para esse grupo. O valor de Z é de 5,123 que reforça a existência da diferença entre as médias, porém de forma não tão substancial quanto no grupo homens com conhecimento.

Figura 26 - Resultado da ANOVA corrigido para grupo mulheres com conhecimento.

Testes de efeitos entre assuntos

Medir: MEASURE_1

Origem		Tipo III Soma dos Quadrados	df	Quadrado Médio	Z	Sig.
Rodadas	Esfericidade considerada	20719,957	6	3453,326	5,123	,001
	Greenhouse-Geisser	20719,957	1,343	15423,315	5,123	,041

Fonte: O autor (2024).

Outro aspecto importante a ser analisado, é a forma que a variação dos tempos entre as rodadas pode ser descrita, essa descrição é feita através do teste de contraste entre as rodadas. A Tabela 17 apresenta o resultado dos testes de contraste para cada um dos seis modelos propostos, analisando o p-valor de cada um dos modelos podemos concluir que os modelos linear, quadrático, cúbico e de ordem4, foram estatisticamente significantes. Através do valor de Z conseguimos ordenar qual dos modelos se adequa melhor a variação das observações, a ordem e dada por linear, cúbica, quadrática e de ordem 4.

Tabela 17 - Teste de contraste para o grupo mulheres com conhecimento

Testes de contrastes entre assuntos

Medir: MEASURE_1

Origem	rodadas	Tipo III Soma dos Quadrados	df	Quadrado Médio	Z	Sig.
rodadas	Linear	22149,348	1	22149,348	31,196	,000
	Quadrático	6982,852	1	6982,852	16,986	,001
	Cúbico	2521,127	1	2521,127	23,645	,000
	Ordem 4	6844,040	1	6844,040	10,397	,007
	Ordem 5	282,064	1	282,064	,680	,426
	Ordem 6	591,798	1	591,798	1,934	,190
Erro(rodadas)	Linear	8520,161	12	710,013		
	Quadrático	4933,185	12	411,099		
	Cúbico	1279,502	12	106,625		
	Ordem 4	7899,431	12	658,286		
	Ordem 5	4974,333	12	414,528		
	Ordem 6	3671,947	12	305,996		

Fonte: O autor (2024).

A pergunta que surgiu ao longo das análises do grupo homens com conhecimento, reaparece nesse grupo e será respondida através do método de Sadak para analisar todos os pares de rodadas possíveis. A Tabela 18 apresenta os resultados das comparações entre os pares de rodadas, ao confrontar o p-valor das comparações par a par com o nível de significância de 0,05, chegamos à conclusão de que apenas a primeira rodada tem uma média significativamente diferente quando comparado com as outras rodadas. Ao comparar a Figura 17, que mostra o tempo médio de duração de cada rodada para o grupo mulheres com conhecimento, vemos que o comportamento do gráfico é mais suave que o grupo homens com conhecimento, apesar de existir uma diferença aparentemente significativa na sexta rodada a comparação pareada comprova que não existe uma diferença significativa da sexta rodada quando comparada com as rodadas 2, 3, 4, 5 e 7.

Tabela 18 - Teste pós-hoc para o grupo mulheres com conhecimento.

Comparações de pares

Medir: MEASURE_1

(I) rodadas	(J) rodadas	Diferença média (I-J)	Erro Padrão	Sig. ^b	95% Intervalo de Confiança para Diferença ^b	
					Limite inferior	Limite superior
1	2	54,399 [*]	11,053	,007	12,175	96,623
	3	45,759 [*]	8,473	,003	13,392	78,126
	4	51,468 [*]	10,285	,006	12,181	90,756

	5	55,072*	11,592	,010	10,790	99,354
	6	70,165*	11,080	,001	27,842	112,489
	7	59,191*	10,193	,002	20,254	98,128
2	1	-54,399*	11,053	,007	-96,623	-12,175
	3	-8,640	7,686	,999	-38,001	20,721
	4	-2,931	10,333	1,000	-42,402	36,541
	5	,673	5,604	1,000	-20,733	22,079
	6	15,766	6,358	,461	-8,523	40,055
	7	4,792	7,763	1,000	-24,862	34,445
3	1	-45,759*	8,473	,003	-78,126	-13,392
	2	8,640	7,686	,999	-20,721	38,001
	4	5,709	5,354	1,000	-14,745	26,163
	5	9,313	6,715	,988	-16,336	34,963
	6	24,406	6,995	,090	-2,315	51,127
	7	13,432	5,957	,608	-9,326	36,189
4	1	-51,468*	10,285	,006	-90,756	-12,181
	2	2,931	10,333	1,000	-36,541	42,402
	3	-5,709	5,354	1,000	-26,163	14,745
	5	3,604	9,325	1,000	-32,018	39,225
	6	18,697	7,381	,428	-9,500	46,894
	7	7,722	6,207	,997	-15,989	31,434
5	1	-55,072*	11,592	,010	-99,354	-10,790
	2	-,673	5,604	1,000	-22,079	20,733
	3	-9,313	6,715	,988	-34,963	16,336
	4	-3,604	9,325	1,000	-39,225	32,018
	6	15,093	7,389	,749	-13,134	43,320
	7	4,118	5,978	1,000	-18,716	26,953
6	1	-70,165*	11,080	,001	-112,489	-27,842
	2	-15,766	6,358	,461	-40,055	8,523
	3	-24,406	6,995	,090	-51,127	2,315
	4	-18,697	7,381	,428	-46,894	9,500
	5	-15,093	7,389	,749	-43,320	13,134
	7	-10,975	4,041	,328	-26,410	4,461
7	1	-59,191*	10,193	,002	-98,128	-20,254
	2	-4,792	7,763	1,000	-34,445	24,862
	3	-13,432	5,957	,608	-36,189	9,326
	4	-7,722	6,207	,997	-31,434	15,989
	5	-4,118	5,978	1,000	-26,953	18,716
	6	10,975	4,041	,328	-4,461	26,410

Fonte: O autor (2024).

4.1.2.3. Grupo homens sem conhecimento

Para o grupo homens sem conhecimento também é necessário realizar o teste de esfericidade de Mauchly. A Figura 27 apresenta o resultado do teste de esfericidade, o p-valor encontrado foi de 0,07 que é maior que o nível de significância de 0,05, constatando que o pressuposto da esfericidade é obedecido.

Figura 27 - Teste de esfericidade para o grupo homens sem conhecimento

Teste de esfericidade de Mauchly^a

Medir: MEASURE_1

Efeito entre assuntos	W de Mauchly	Aprox. Qui-quadrado	df	Sig.	Epsilon ^b		
					Greenhouse-Geisser	Huynh-Feldt	Limite inferior
Rodadas	,070	31,333	20	,058	,494	,641	,167

Fonte: o autor (2024).

Após a verificação da esfericidade, é possível aplicar a ANOVA com medidas repetidas, para avaliar a existência de diferença significativa da média entre as rodadas. A Figura 28 apresenta o resultado da análise de variância, comparando o p-valor encontrado com o valor do nível de significância, concluímos que existe uma diferença estatisticamente significativa entre as médias de cada rodada. E o alto valor de Z corrobora para a conclusão da diferença estatística entre as médias, além disso o valor de 15,564 constata que a diferença das médias é ainda mais substancial que a diferença encontrada no grupo homens com conhecimento.

Figura 28 - Resultado da ANOVA para grupo homens sem conhecimento.

Testes de efeitos entre assuntos

Medir: MEASURE_1

Origem		Tipo III Soma dos Quadrados	df	Quadrado Médio	Z	Sig.
Rodadas	Esfericidade considerada	24543,526	6	4090,588	15,564	,000

Fonte: O autor (2024).

A constatação da existência de uma diferença significativa entre as médias do grupo de homens sem conhecimento, levanta a discussão de como as variâncias entre as observações das rodadas podem ser representadas. Avaliando a Tabela 19, identificamos que os modelos que melhor representam essa variação pode ser linear, quadrático e de ordem 4. E analisando o valor

de Z temos que o modelo que mais representa as variações são o linear, seguido do de ordem 4 e por último quadrático.

Tabela 19 - Teste de contraste para o grupo homens sem conhecimento

Testes de contrastes entre assuntos

Medir: MEASURE_1

Origem	Rodadas	Tipo III Soma dos Quadrados	df	Quadrado Médio	Z	Sig.
Rodadas	Linear	11979,436	1	11979,436	58,334	,000
	Quadrático	3766,108	1	3766,108	6,431	,024
	Cúbico	715,547	1	715,547	2,701	,123
	Ordem 4	7887,375	1	7887,375	33,770	,000
	Ordem 5	178,066	1	178,066	1,638	,221
	Ordem 6	16,995	1	16,995	,095	,762
Erro(Rodadas)	Linear	2875,049	14	205,361		
	Quadrático	8198,149	14	585,582		
	Cúbico	3708,829	14	264,916		
	Ordem 4	3269,853	14	233,561		
	Ordem 5	1522,386	14	108,742		
	Ordem 6	2503,122	14	178,794		

Fonte: O autor (2024).

A Tabela 20 mostra a comparação par a par, através do teste de Sadak, dessa maneira confrontamos a significância da comparação com o nível de significância para concluir quais rodadas possuem médias estatisticamente significativa. A primeira e a sexta rodada apresentaram um valor de significância menor que o nível de significância de 0,05, quando comparadas com as outras rodadas, diferindo dos grupos de homens com conhecimento e mulheres com conhecimento, evidenciando que as médias dessas duas rodadas são estatisticamente diferentes. Comparando as análises da Tabela 20 com o gráfico apresentado na Figura 19, vemos que as diferenças visíveis entre a rodada 1 e 6 no gráfico são estatisticamente significativas, o que não ocorreu com os dois grupos anteriores, onde apesar de visivelmente existir uma diferença o teste de Sadak não atestou que essa diferença é significativa.

Tabela 20 - Teste pós-hoc para o grupo homens sem conhecimento.

Comparações de pares

Medir: MEASURE_1

(I) rodadas	(J) rodadas	Diferença média (I-J)	Erro Padrão	Sig. ^b	95% Intervalo de Confiança para	
					Diferença ^b	
					Limite inferior	Limite superior
1	2	36,468 [*]	9,460	,036	1,588	71,348
	3	34,861 [*]	8,159	,016	4,779	64,943
	4	29,217	8,285	,068	-1,329	59,764
	5	36,143 [*]	7,724	,007	7,665	64,622
	6	55,199 [*]	7,675	,000	26,900	83,498
	7	36,931 [*]	6,339	,001	13,558	60,304
	2	1	-36,468 [*]	9,460	,036	-71,348
3		-1,607	4,597	1,000	-18,556	15,341
4		-7,251	4,675	,961	-24,489	9,987
5		-,325	5,244	1,000	-19,658	19,009
6		18,731 [*]	4,813	,034	,985	36,476
7		,463	6,240	1,000	-22,543	23,469
3		1	-34,861 [*]	8,159	,016	-64,943
	2	1,607	4,597	1,000	-15,341	18,556
	4	-5,643	4,935	,999	-23,839	12,553
	5	1,283	3,721	1,000	-12,438	15,003
	6	20,338 [*]	4,385	,008	4,170	36,506
	7	2,071	4,794	1,000	-15,605	19,747
	4	1	-29,217	8,285	,068	-59,764
2		7,251	4,675	,961	-9,987	24,489
3		5,643	4,935	,999	-12,553	23,839
5		6,926	3,643	,819	-6,507	20,359
6		25,981 [*]	4,348	,001	9,950	42,012
7		7,714	6,549	,998	-16,433	31,861
5		1	-36,143 [*]	7,724	,007	-64,622
	2	,325	5,244	1,000	-19,009	19,658
	3	-1,283	3,721	1,000	-15,003	12,438
	4	-6,926	3,643	,819	-20,359	6,507
	6	19,055 [*]	4,288	,012	3,247	34,863
	7	,788	5,107	1,000	-18,042	19,618
	6	1	-55,199 [*]	7,675	,000	-83,498
2		-18,731 [*]	4,813	,034	-36,476	-,985
3		-20,338 [*]	4,385	,008	-36,506	-4,170
4		-25,981 [*]	4,348	,001	-42,012	-9,950

	5	-19,055*	4,288	,012	-34,863	-3,247
	7	-18,267*	4,460	,023	-34,712	-1,823
7	1	-36,931*	6,339	,001	-60,304	-13,558
	2	-,463	6,240	1,000	-23,469	22,543
	3	-2,071	4,794	1,000	-19,747	15,605
	4	-7,714	6,549	,998	-31,861	16,433
	5	-,788	5,107	1,000	-19,618	18,042
	6	18,267*	4,460	,023	1,823	34,712

Fonte: O autor (2024).

4.1.2.4. Grupo mulheres sem conhecimento

A Figura 29 traz o teste de esfericidade para o grupo mulheres sem conhecimento. O p-valor do teste de Mauchly foi menor que 0,001, sugerindo que o pressuposto da esfericidade foi violado. E que a correção pelo método de Greenhouse-Geisser é necessária para realizar o teste da ANOVA com medidas repetidas.

Figura 29 - Teste de esfericidade para o grupo mulheres sem conhecimento

Teste de esfericidade de Mauchly^a

Medir: MEASURE_1

Efeito entre assuntos	W de Mauchly	Aprox. Qui-quadrado	df	Sig.	Epsilon ^b		
					Greenhouse-Geisser	Huynh-Feldt	Limite inferior
Rodadas	,000	.	20	.	,325	,530	,167

Fonte: O autor (2024).

O teste de análise de variância corrigido pelo fator de Greenhouse-geisser, mostra que não existe diferença entre as médias das rodadas para esse grupo. Tendo como base a comparação entre a significância do teste com o nível de significância de 0,05, como o valor foi superior ao nível de significância comprova a não existência de diferenças entre as médias. O baixo valor de Z também indica essa indiferença, além disso a visualização do gráfico apresentado na Figura 21 sugere que a rodada 1, 4 e 5 tem uma diferença acentuada na média, porém a ANOVA revela que essa diferença é estatisticamente insignificante.

Figura 30 - Resultado da ANOVA para grupo mulheres sem conhecimento.

Testes de efeitos entre assuntos

Medir: MEASURE_1

Origem		Tipo III Soma dos Quadrados	df	Quadrado Médio	Z	Sig.
Rodadas	Esfericidade considerada	4167,943	6	694,657	2,487	,045
	Greenhouse-Geisser	4167,943	1,949	2138,539	2,487	,135

Fonte: O autor (2024).

Apesar de não existir uma diferença significativa, ainda sim é importante analisar como as variações entre as rodadas podem ser descritas. A Tabela 21 apresenta o resultado do teste de contraste que analise as principais formas de descrever a variação, o grupo de mulheres sem conhecimento apresenta amenas dois modelos que descrevem bem as variações entre as rodadas o modelo linear e o quadrático, onde o modelo linear descreve melhor que o modelo quadrático.

Tabela 21 - Teste de contraste para o grupo mulheres sem conhecimento

Testes de contrastes entre assuntos

Medir: MEASURE_1

Origem	Rodadas	Tipo III Soma dos Quadrados	df	Quadrado Médio	Z	Sig.
Rodadas	Linear	1942,488	1	1942,488	14,383	,013
	Quadrático	1102,719	1	1102,719	7,279	,043
	Cúbico	649,995	1	649,995	,775	,419
	Ordem 4	76,616	1	76,616	,200	,674
	Ordem 5	29,697	1	29,697	,275	,622
	Ordem 6	366,429	1	366,429	6,195	,055
Erro(Rodadas)	Linear	675,271	5	135,054		
	Quadrático	757,505	5	151,501		
	Cúbico	4191,660	5	838,332		
	Ordem 4	1919,300	5	383,860		
	Ordem 5	539,395	5	107,879		
	Ordem 6	295,727	5	59,145		

Fonte: O autor (2024).

4.1.2.5. Comparação entre os grupos

Até esse momento foram comparadas as médias de um mesmo grupo ao longo das sete etapas que envolvem a avaliação intracritério. Nesse momento se torna importante a avaliação das médias entre os diferentes grupos. Para realizar essa comparação será usado a metodologia proposta por Tukey (1953). O teste de turkey, também conhecido como Honestly Significant Difference (HSD), é um mecanismo para realizar comparações múltiplas entre diferentes grupos, com o objetivo de identificar a diferença estatística entre os grupos. O teste atua controlando o erro do tipo I reduzindo a incidência de falsos positivos, partindo do ajuste dos valores de p após uma análise de variância. O Quadro 3 apresenta o resultado do teste de Turkey para os quatro grupos definidos. Ao analisar as comparações entre os grupos notamos que todos os valores de p são maiores que os nível de significância de 0,05, sugerindo que não existe uma diferença significativa entre os quatro grupos.

Quadro 3 - Teste de Turkey para comparação entre os grupos.

Multiple Comparison of Means - Tukey HSD, FWER=0.05

group1	group2	meandiff	p-adj	lower	upper	reject
Homens com conhecimento	Homens sem conhecimento	-4.2771	0.9706	-30.9806	22.4263	False
Homens com conhecimento	Mulheres com conhecimento	-8.81	0.7996	-35.5134	17.8934	False
Homens com conhecimento	Mulheres sem conhecimento	-12.7929	0.5587	-39.4963	13.9106	False
Homens sem conhecimento	Mulheres com conhecimento	-4.5329	0.9653	-31.2363	22.1706	False
Homens sem conhecimento	Mulheres sem conhecimento	-8.5157	0.8153	-35.2191	18.1877	False
Mulheres com conhecimento	Mulheres sem conhecimento	-3.9829	0.976	-30.6863	22.7206	False

Fonte: O autor (2024).

Ainda analisando o Quadro 3, é possível abordar os intervalos de confiança que a diferença da média entre as comparações está inserida. Para todas as comparações temos que o limite inferior (lower) e o limite superior (upper), dos intervalos de confiança, estão compreendidos entre um número negativo (limite inferior) e um número positivo (limite superior) essa característica reforça a não existência de diferença significativa entre os grupos. Essa suposição parte do princípio de que o intervalo de confiança que a verdadeira diferença da média está contida, é calculado como uma faixa de valores para um nível de confiança de 95%, dessa maneira quando o zero é incluído não se pode afirmar que a média entre os grupos diferem, pois, a diferença pode ser zero.

Contudo, apesar do teste de Turkey revelar que não existe uma diferença entre as médias de cada grupo, se torna importante avaliar se existe uma diferença entre os grupos rodada a rodada. A Figura 31 apresenta o teste de kruskal-wallis, o teste avalia todos os grupos e investiga a existência de uma diferença estatisticamente significativa. Analisando a figura, é

possível identificar que existe uma diferença significativa entre os grupos apenas na quarta rodada

Figura 31 - Teste estatístico para comparar as rodadas.

	r1	r2	r3	r4	r5	r6	r7
Qui-quadrado	4,458	,895	4,155	6,827	2,133	5,440	2,576
df	3	3	3	3	3	3	3
Significância Sig.	,216	,827	,245	,078	,545	,142	,462

a. Teste Kruskal Wallis

b. Variável de Agrupamento: Grupos

Fonte: O autor (2024).

4.1.3. Análise de correlação

Na sessão anterior foi comprovado que não existe uma diferença significativa entre os grupos, contudo analisando os grupos de forma individual observamos que poucas rodadas têm uma diferença entre as médias que fossem estatisticamente significativas. Apesar disso, os gráficos das medias dos tempos de execução de cada rodada, apresenta uma tendência de redução no tempo que os participantes levam para concluir cada etapa da avaliação.

Essa pesquisa sugere que a causa para esse efeito é proveniente de duas hipóteses. A primeira hipótese, parte da relação e do efeito que uma rodada tem sobre a outra, levando a tendência decrescente que foi citada. A segunda hipótese, é que essa tendência decrescente é causada pelo efeito do aprendizado dos participantes sobre o experimento. Essa sessão irá investigar o efeito e a causa que uma rodada tem sobre a outra.

A análise de correlação é uma técnica que se utiliza de estatística para avaliar a relação entre duas ou mais variáveis, de forma que compreender a força e a direção que essa relação tem, porém sem criar qualquer implicação de causalidade. Os dois métodos mais populares para analisar a correlação entre variáveis, foram sugeridos por Pearson (1896) e por Spearman (1904). O teste de Pearson é utilizado com os dados seguem uma distribuição normal enquanto o teste de Spearman não requer uma normalidade das observações, dessa maneira será necessário a avaliar o pressuposto de normalidade para toda a amostra, e partindo do resultado aplicar a correlação mais adequada.

Tabela 22 - Teste de normalidade para amostra.

Rodada	Estatística	Valor-p	Normalidade
1ª	0.9236	0.0079	Não é normal
2ª	0.9556	0.1027	É normal
3ª	0.9214	0.0067	Não é normal
4ª	0.9455	0.0447	Não é normal
5ª	0.8772	0.0003	Não é normal
6ª	0.6779	0.0000	Não é normal
7ª	0.9695	0.3176	É normal

Fonte: O autor (2024).

A Tabela 22 apresenta os resultados dos testes de Shapiro-wilk para avaliar a normalidade das rodadas para toda a amostra. Analisando a tabela é possível concluir que apenas duas das sete rodadas apresenta normalidade, desse modo o teste que mais se adequa aos dados é o teste de Spearman. A Tabela 23 mostra a matriz de correlação de Spearman, em que, valores próximos a 1 representa uma correlação positiva, valores próximos a -1 representam uma correlação negativa e valores próximos a 0 apresenta relações fracas (positivas ou negativas).

Tabela 23 - Matriz de correlação para amostra.

Rodadas	Rodada 1	Rodada 2	Rodada 3	Rodada 4	Rodada 5	Rodada 6	Rodada 7
Rodada 1	1.000.000	0.329311	0.547334	0.303398	0.264148	0.380911	0.357445
Rodada 2	0.329311	1.000.000	0.354847	0.011794	0.363247	0.242209	0.093903
Rodada 3	0.547334	0.354847	1.000.000	0.603852	0.562112	0.392774	0.496761
Rodada 4	0.303398	0.011794	0.603852	1.000.000	0.420006	0.302489	0.432606
Rodada 5	0.264148	0.363247	0.562112	0.420006	1.000.000	0.502003	0.562410
Rodada 6	0.380911	0.242209	0.392774	0.302489	0.502003	1.000.000	0.481867
Rodada 7	0.357445	0.093903	0.496761	0.432606	0.562410	0.481867	1.000.000

Fonte: O autor (2024).

A Tabela 24, apresenta os valores da correlação e o tipo de relação associado. Os valores da diagonal principal apresentam uma correlação perfeita, já que ela compara uma rodada com ela mesma. Os valores fora da diagonal principal apresentam no máximo uma relação moderada, com destaque para a relação moderada entre as rodadas 3 e 4 (coeficiente de 0,60) e para a relação entre as rodadas 1 e 3 (coeficiente de 0,54). Já as relações muito fracas têm como destaque a relação entre as rodadas 2 e 4 (coeficiente de 0,011) e entre as rodadas 2 e 7 (coeficiente 0,093). A Tabela 25 mostra a melhor e a pior relação que cada rodada tem com as outras, vemos que a rodada 5 aparece três vezes como a melhor relação e uma vez como pior

relação, já a rodada 2 aparece quatro vezes como pior relação e não aparece como melhor relação em nenhuma das outras rodadas.

Tabela 24 - Descrição dos valores de correlação.

Valor da correlação positiva	Valor da correlação Negativa	Tipo da relação
Entre 0,8 e 1	Entre -0,8 e -1	Muito forte
Entre 0,6 e 0,8	Entre -0,6 e -0,8	Forte
Entre 0,4 e 0,6	Entre -0,4 e -0,6	Moderada
Entre 0,2 e 0,4	Entre -0,2 e -0,4	Fraca
Entre 0 e 0,2	Entre -0 e -0,2	Muito fraca

Fonte: O autor (2024).

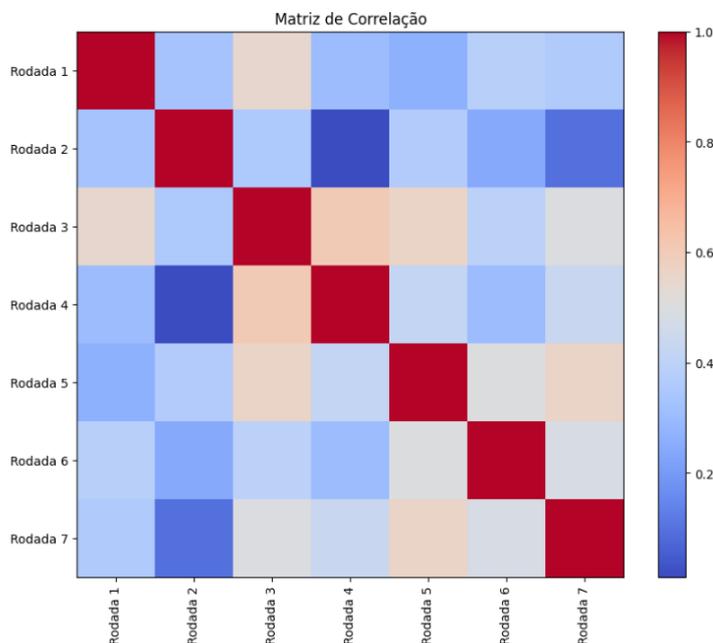
Tabela 25 - Melhor e pior relação entre as rodadas.

Rodadas	Melhor relação / Coeficiente	Pior relação / Coeficiente
Rodada 1	Rodada 3 / 0,54	Rodada 5 / 0,26
Rodada 2	Rodada 5 / 0,36	Rodada 4 / 0,01
Rodada 3	Rodada 4 / 0,60	Rodada 2 / 0,35
Rodada 4	Rodada 3 / 0,60	Rodada 2 / 0,01
Rodada 5	Rodada 7 / 0,56	Rodada 1 / 0,26
Rodada 6	Rodada 5 / 0,50	Rodada 2 / 0,24
Rodada 7	Rodada 5 / 0,56	Rodada 2 / 0,09

Fonte: O autor (2024).

A Figura 32 traz uma visualização alternativa da matriz de correlação. Ao avaliar a matriz de correlação de forma mais profunda, percebemos que 57% das relações entre as variáveis são classificadas como fraca ou muito fraca e 43% das relações são classificadas como moderada. Tendo essa avaliação como base podemos concluir que as rodadas têm pouca influência no desempenho dos participantes.

Figura 32 - Visualização gráfica da matriz de correlação.



Fonte: O autor (2024).

As análises da matriz de correlação de Spearman, sugerem que as relações entre as rodadas são em sua maioria fraca ou muito fraca, contudo é preciso concluir se apesar de fraca essas relações são significantes. Gosset (1908) introduz o conceito da estatística t, a distribuição t é especialmente eficaz para um conjunto de dados pequenos e que seguem uma distribuição normal, e tem como particularidade a irrelevância do conhecimento da variância populacional. Tabela 22 comprovou que a amostra não segue uma distribuição normal para algumas rodadas, contudo Rice (2007) e Casella e Berger (2002) contextualizam que apesar da distribuição t ter o pressuposto da normalidade dos dados, o princípio do teorema do limite central (TLC) justifica que a distribuição t pode ser utilizada em dados que violão a normalidade, desde que o tamanho da amostra seja grande. O TLC justifica essa aplicação afirmando que para amostras de tamanho grande, quando a amostra tem mais de 30 observações, as distribuições não normais se aproximam de uma distribuição normal.

DeGroot e Schervish (2012) dizem que o teste t é um poderoso método para testar hipóteses e ela é amplamente usada para determinar valores críticos para testes de hipótese. A Tabela 26 apresenta o valor da estatística t que estão relacionados aos valores da correlação de Spearman mostrado na Tabela 23.

Tabela 26 - Valor da estatística t.

Rodadas	Rodada 1	Rodada 2	Rodada 3	Rodada 4	Rodada 5	Rodada 6	Rodada 7
Rodada 1	-	2.206	4.136	2.014	1.732	2.606	2.421
Rodada 2	2.206	-	2.400	0.075	2.466	1.579	0.597
Rodada 3	4.136	2.400	-	4.791	4.298	2.701	3.620
Rodada 4	2.014	0.075	4.791	-	2.927	2.007	3.035
Rodada 5	1.732	2.466	4.298	2.927	-	3.671	4.302
Rodada 6	2.606	1.579	2.701	2.007	3.671	-	3.478
Rodada 7	2.421	0.597	3.620	3.035	4.302	3.478	-

Fonte: O autor (2024).

Antes de prosseguir, é preciso definir o teste de hipótese a qual será submetido as estatísticas t apresentada anteriormente. A hipótese nula, H_0 , é definida como a correlação entre as rodadas não é significativa, já a hipótese alternativa, H_1 , diz que a correlação entre as rodadas é significativa. A Tabela 27, traz o valor de P associado a cada estatística t apresentada na Tabela 26. O valor de p é comparado com o nível de confiança de 0,10, nos casos que o p-valor é menor que o nível de significância rejeitamos H_0 , e concluímos que a correlação é significativa, por outro lado se o p-valor for maior que 0,10, deixamos de rejeitar H_0 , concluindo que a correlação não tem significância estatística. Ao analisar a Tabela 27, é possível ver que apenas 6 dos 42 valores, foram acima de 0,10, dando um total de 86% de valores significativos e 14% de valores não significativos.

Tabela 27 - p-valor associado a estatística t

Rodada	Rodada 1	Rodada 2	Rodada 3	Rodada 4	Rodada 5	Rodada 6	Rodada 7
Rodada 1	1	0.033	0.000	0.051	0.091	0.013	0.020
Rodada 2	0.033	1	0.021	0.941	0.018	0.122	0.554
Rodada 3	0.000	0.021	1	0.000	0.000	0.010	0.001
Rodada 4	0.051	0.941	0.000	1	0.006	0.052	0.004
Rodada 5	0.091	0.018	0.000	0.006	1	0.001	0.000
Rodada 6	0.013	0.122	0.010	0.052	0.001	1	0.001
Rodada 7	0.020	0.554	0.001	0.004	0.000	0.001	1

Fonte: O autor (2024).

Analisando de forma conjunta o teste t e a correlação de Spearman, temos as evidências necessárias para dizer que, as rodadas influenciam na tendencia decrescente no tempo de avaliação de cada critério, na etapa intracritério do FITradeoff. Apesar de existir uma comprovação de que existe uma relação estatisticamente significativa do efeito que uma rodada tem sobre a rodadas, a influência é sutil e em alguns casos pode chegar a ser moderada, dito

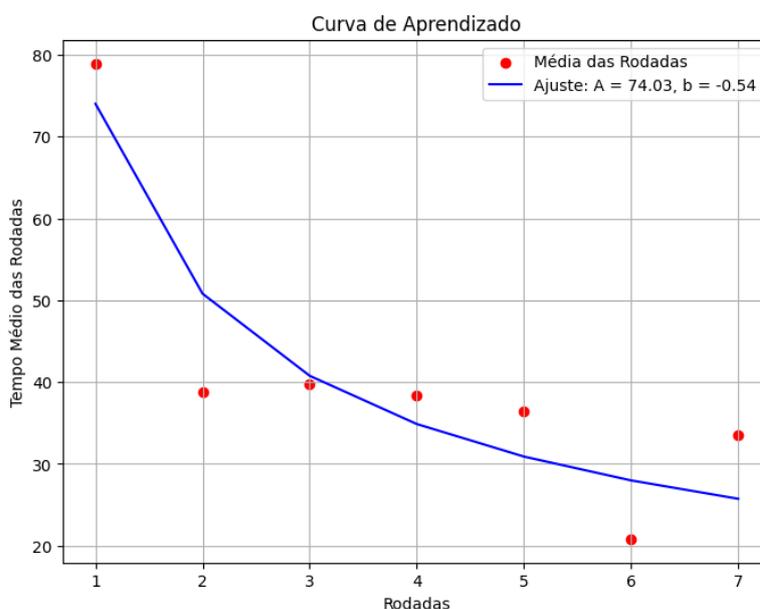
isto, concluímos que o efeito das rodadas ao longo do experimento, influencia no decaimento no tempo de execução, porém de forma modéstia.

4.1.4. Curva de aprendizagem

Essa sessão irá investigar a segunda hipótese, na qual sugere que a tendência de decréscimo dos tempos de execução das rodadas é causada pelo aprendizado dos participantes ao longo do experimento.

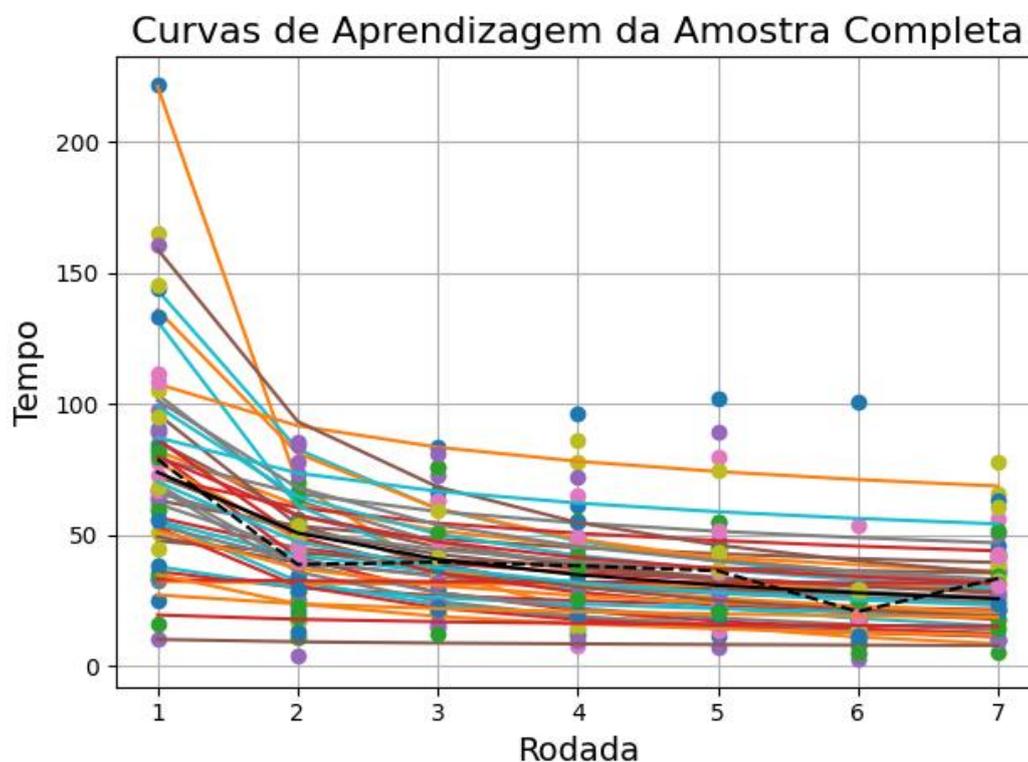
A curva de aprendizagem será traçada com base no modelo descrito pela equação 1. Esse modelo tem dois parâmetros A e o B, o parâmetro A representa o desempenho do participante na primeira rodada, já o parâmetro B define a taxa de aprendizagem dos participantes. Partindo do parâmetro B é possível descobrir a porcentagem de aprendizado p , que é definida como 2^{-b} . Para traçar a curva de aprendizagem é preciso calcular os parâmetros, segundo Montgomery *et al.* (2021), existem duas formas de estimar o valor dos parâmetros, a primeira maneira é através de regressão linear e a segunda por meio da transformação logarítmica, esse trabalho utilizará o método da transformação logarítmica. A Figura 33 mostra a curva de aprendizagem traçada para a amostra e os parâmetros estimados. Analisado a imagem, vemos que o valor de A e B são, respectivamente, 74,03 e -0,54. A Figura 34, exibe a curva de aprendizado de todos os participantes que compõem a amostra.

Figura 33 - Curva de aprendizagem a amostra.



Fonte: O autor (2024).

Figura 34 - Curva de aprendizagem da amostra completa



Fonte: O autor (2024).

Analisando a curva de aprendizagem mostrada na Figura 33, notamos que a taxa de aprendizagem geral da amostra é de $-0,54$. O sinal negativo do parâmetro B indica que existe uma redução no tempo da atividade causada pelo efeito do conhecimento. Por outro lado, o valor de B, indicado que o decrescimento é mais acentuado nas primeiras rodadas e vai diminuindo até se estabilizar. O comportamento assintótico desse modelo, garante que com o aumento do número de repetições o tempo de execução nunca chegará a zero, além de intuir que a melhoria no desempenho acontecerá nas rodadas iniciais, constatando que existe uma desaceleração da aprendizagem.

Comparando as ideias descritas no parágrafo anterior e os gráficos dos tempos médios de execução para os grupos definidos. O comportamento assintótico se torna evidente, independente do gráfico analisado, que o tempo da primeira rodada é significativamente maior que o tempo das outras e que as rodadas finais têm uma consistência maior que as iniciais.

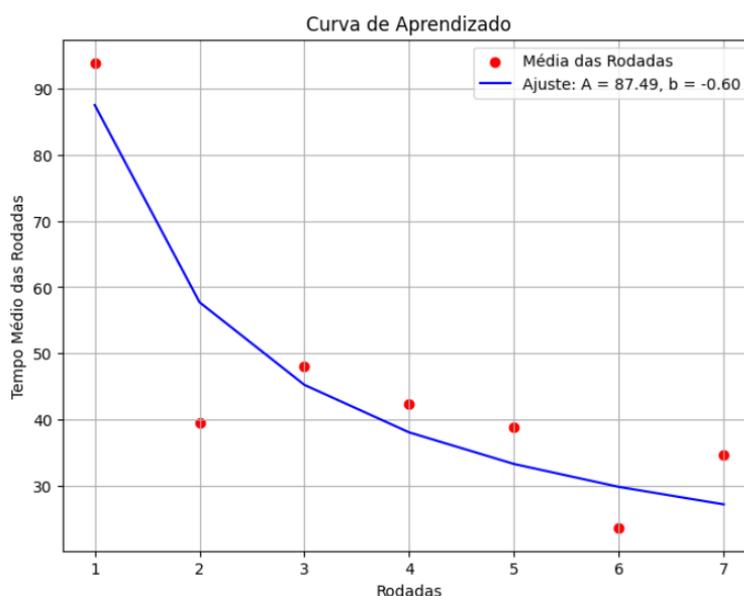
Outro discurso pertinente, é acerca da sexta rodada que para todos os grupos, exceto para o grupo mulheres sem conhecimento, teve o menor tempo de duração. Essa diferença pode estar relacionada as particularidades do experimento e não necessariamente a um aprendizado súbito.

As próximas subseções irão discutir a curva de aprendizado de cada, bem como a avaliação das taxas de aprendizagem e a comparação entre os grupos. Tais discursões são necessárias para entender o relacionamento do conhecimento e do gênero sobre a forma que a aprendizagem influencia na tendência de decrescimento.

4.1.4.1. Grupo homens com conhecimento

Para traçar a curva de aprendizagem, para o grupo homens com conhecimento, foi utilizado os mesmos métodos apresentados na curva da amostra. A Figura 35, mostra a curva de aprendizagem ajustada para o grupo homens com conhecimento, analisando a figura é possível perceber que a curva se ajusta bem aos dados dos tempos médios de cada rodada. Comparando o valor do parâmetro B, da curva de aprendizagem da amostra e do grupo homens com conhecimento, é possível concluir que a curva desse grupo possui um desaceleramento mais acentuado que a curva da amostra, indicando que existe uma dificuldade para continuar aprendendo ao longo do tempo. A taxa de aprendizagem para o grupo homens com conhecimento, alcança o ponto de estagnação mais rápido e de forma mais relevante, consequentemente deixando os participantes em um platô no qual o desempenho na atividade não tem mais melhoria.

Figura 35 - Curva de aprendizagem ajustada grupo homens com conhecimento.

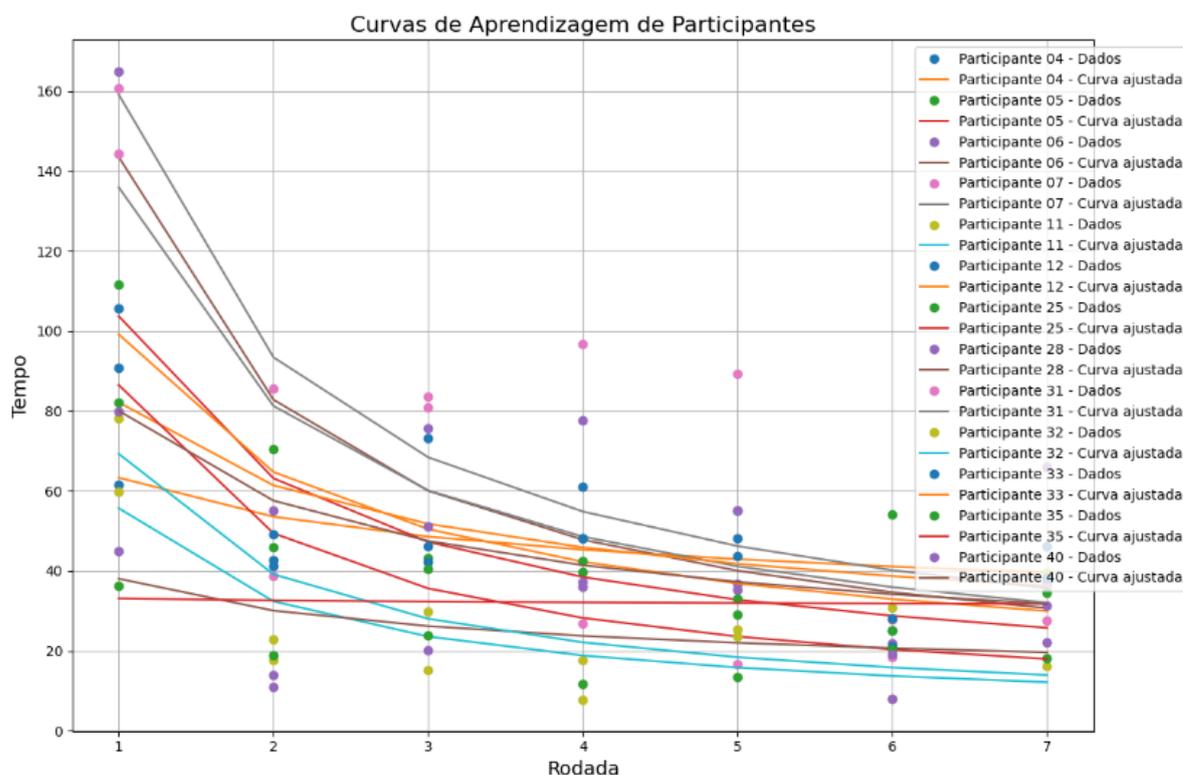


Fonte: O autor (2024).

A Figura 36 mostra a curva de todos os participantes pertencentes a esse grupo. Analisando a figura vemos que a curva de todos os participantes tem comportamento semelhante, porém com suas respectivas particularidades. Ainda analisando a Figura 36, vemos que existe uma alta variedade nas curvas dos participantes, indicando comportamentos como o do participante 32 que apresenta um valor relativamente baixo, no tempo de execução da primeira rodada, enquanto possui um elevado valor de taxa de aprendizagem. Em contrapartida, existe participantes como o participante 12, que tem tempo da primeira rodada aproximadamente igual à do participante 32, mas com metade da taxa de aprendizagem.

De forma geral, os participantes que possuem menor valor absoluto do parâmetro B, conseguem tempos de execução menores e possui curva de aprendizagem mais suaves. Por outro lado, os participantes que possuem valor maior do parâmetro B tendem a ter valores mais elevados na primeira rodada, gerando curvas mais acentuadas e tempos de execução da última rodada maior. Um caso em particular, que vale apenas ser mencionado, é o participante 35, ele possui baixo valor do tempo de execução da primeira rodada e baixo valor da taxa de aprendizagem, contudo seu tempo de execução cresce de forma quase linear, colocando em xeque o comprometimento desse participante com o experimento.

Figura 36 - Curva ajustada dos participantes do grupo homens com conhecimento.

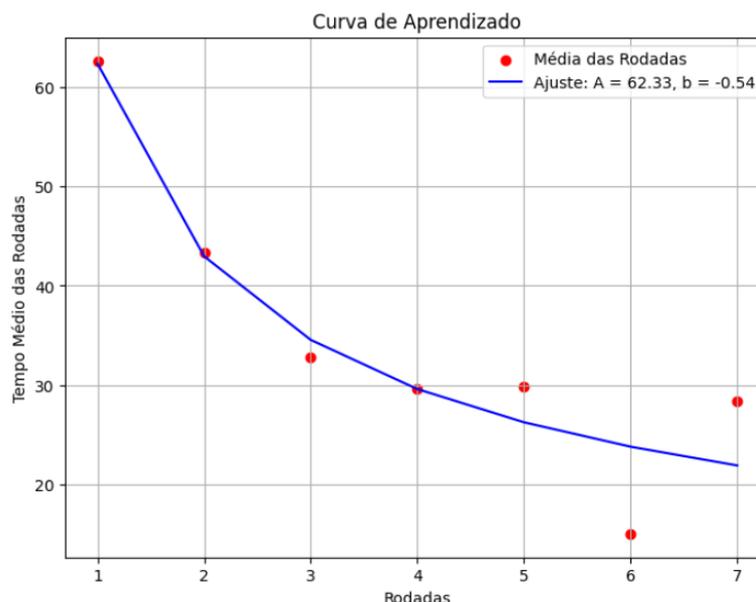


Fonte: O autor (2024).

4.1.4.2. Grupo mulheres com conhecimento

A Figura 37 mostra a curva de aprendizagem para o grupo mulheres com conhecimento. A curva de aprendizagem ajustada, vista na figura, se ajusta bem aos dados, a curva tem parâmetros A e B, respectivamente iguais a, 62,33 e -0,54. Esses parâmetros revelam um comportamento menos acentuado que o grupo homens com conhecimento e a amostra, contudo apesar da taxa de aprendizagem desse grupo ser menor que o grupo, homens com conhecimento, ele também atinge a estabilidade de forma um pouco mais tardia, e não necessariamente esse tempo de estabilidade é mais baixo que os outros.

Figura 37 - Curva de aprendizagem ajustada para o grupo mulheres com conhecimento.

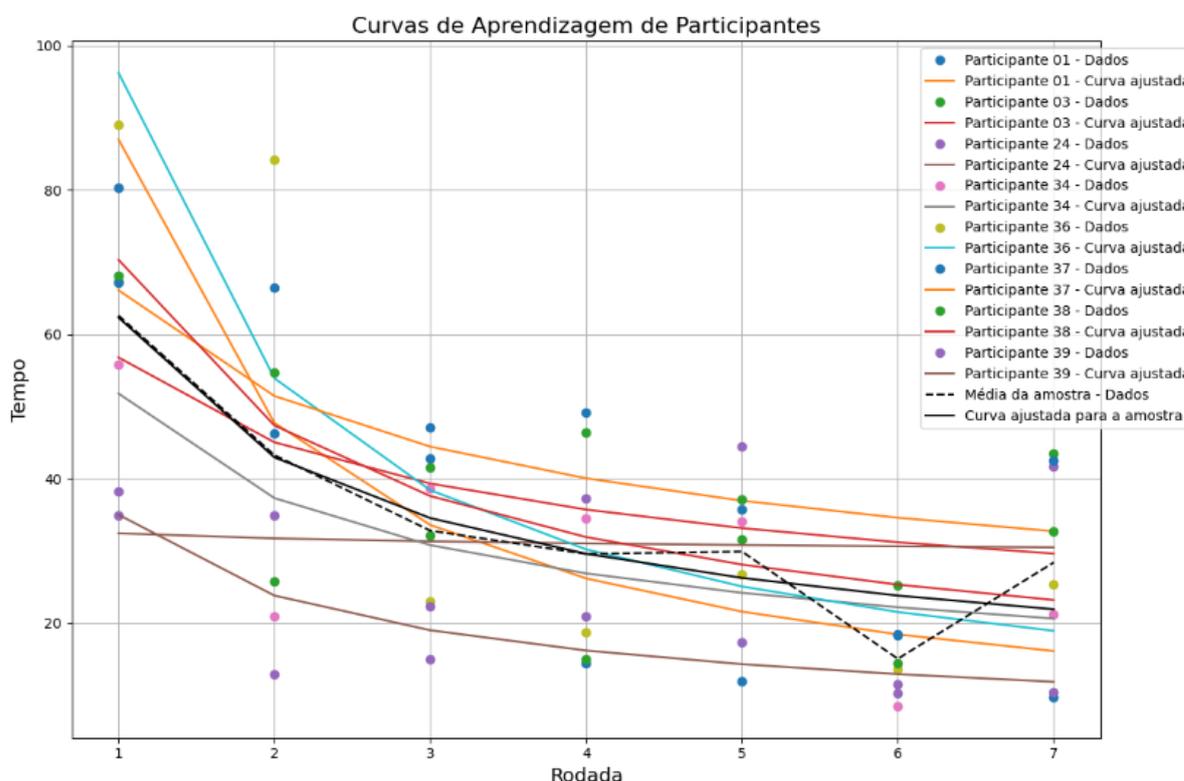


Fonte: O autor (2024).

A Figura 38 mostra as curvas de aprendizagem para os participantes que pertencem a esse grupo. A curva dos participantes, tem menos variação que o grupo homens com conhecimento, as curvas se concentram em uma área específica. Apesar de existir uma variação menor esse grupo possui uma amplitude considerável, com destaque para o participante 01 que tem a maior taxa de aprendizagem e para o participante 24 que possui a menor taxa de aprendizagem, em valores absolutos. Outro aspecto importante, é que diferente do grupo anterior, nesse grupo, os participantes com maior taxa de aprendizagem também tiveram o maior tempo de execução da primeira rodada, por outro lado o participante que obteve a menor taxa de aprendizagem também obteve o menor tempo de duração na primeira rodada, para esse

participante o seu avanço teve um comportamento linear, de forma decrescente, esse comportamento oferece duas suposições distintas, a primeira é que o participante já tinha familiaridade com o SAD FITradeoff a outra é supõe o não comprometimento do participante com o experimento.

Figura 38 - Curva ajustada dos participantes do grupo mulheres com conhecimento.



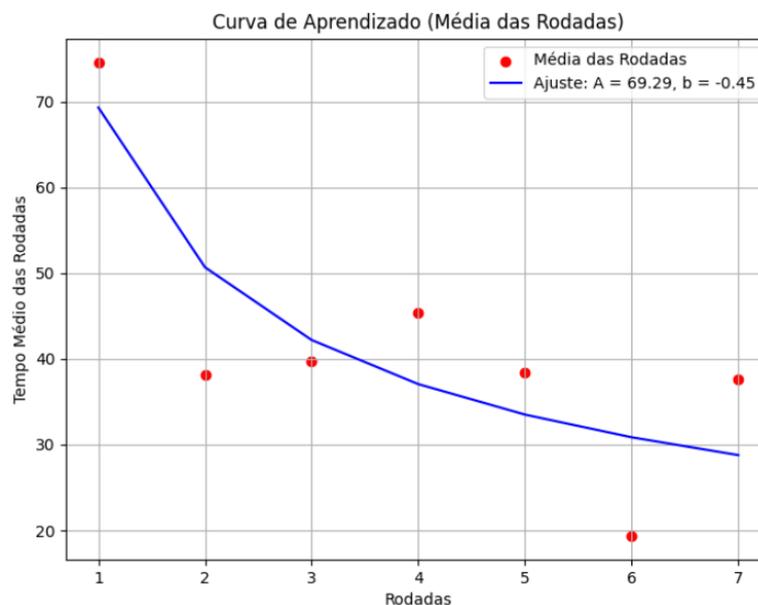
Fonte: O autor (2024).

4.1.4.3. Grupo homens sem conhecimento

A Figura 39 mostra a curva de aprendizagem para o grupo de homens sem conhecimento. A curva ajustada, que pode ser vista na figura, se ajusta bem aos dados das médias do tempo de duração de cada rodada, os parâmetros A e B da curva, são respectivamente 69,29 e -0,45. A aprendizagem desse grupo se comporta de forma decrescente como os outros dois grupos, porém com a menor taxa de aprendizagem dentre os grupos analisados até agora. A curva se mostra mais suave e com uma melhoria mais lenta, em relação aos grupos de homens e mulheres com conhecimento, porém com tempo inicial menor que os homens com

conhecimento. Nesse grupo existe uma melhoria ao longo do tempo, mas quando comparado a curva da amostra o ritmo é mais rápido e o platô de estagnação é açoitado rapidamente.

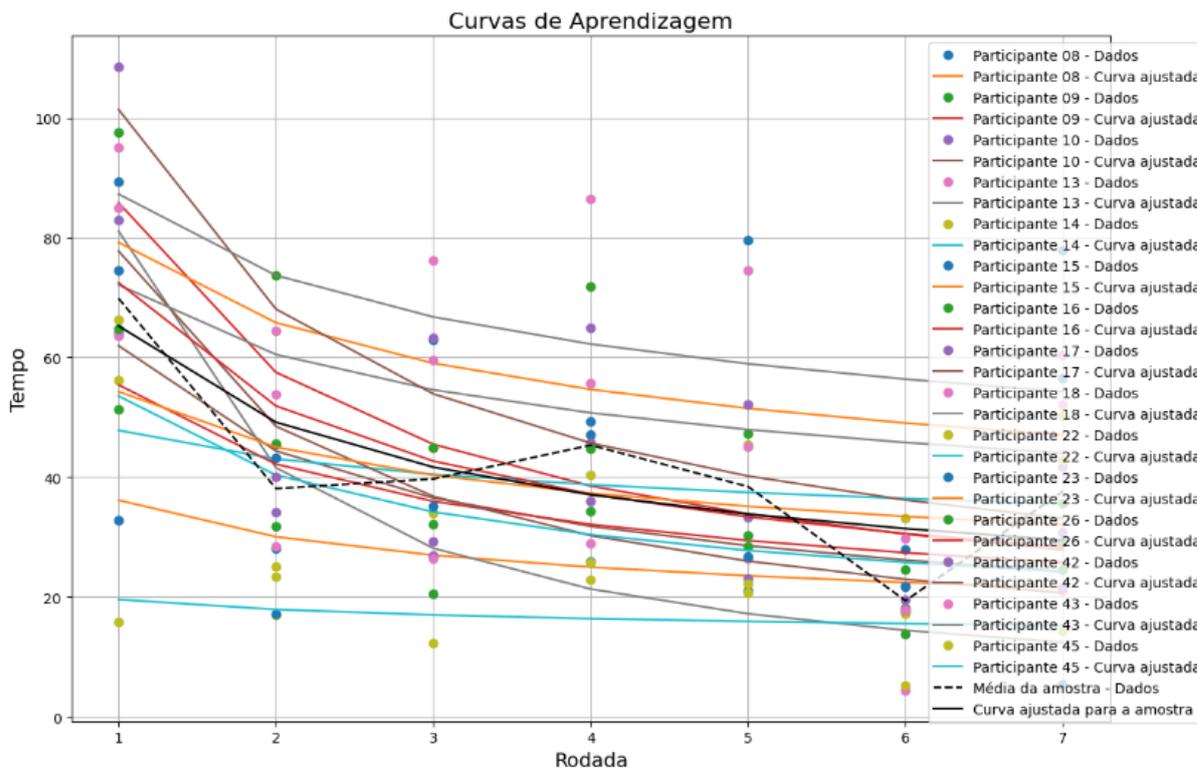
Figura 39 - Curva de aprendizagem ajustada grupo homens sem conhecimento.



Fonte: O autor (2024).

A Figura 40, mostra as curvas de aprendizagem de todos os participantes pertencentes a esse grupo. A variabilidade das curvas desse grupo é semelhante ao do grupo homens com conhecimento, porém com um tempo inicial consideravelmente menor. Esse grupo apresenta uma amplitude relativamente baixa, da taxa de aprendizagem, quando comparado com os outros dois grupos anteriores. Esse grupo como destaque o participante 42 que tem o maior tempo na primeira rodada e um aprendizado moderado, já o participante 17 possui um tempo inicial consideravelmente menor, em comparação com o participante 42, mas sua taxa de aprendizagem é relativamente maior. Esse grupo também apresenta participantes com comportamento linear, sugerindo descomprometimento com o experimento, uma vez que o tempo da primeira rodada é maior que o tempo da última, indo na direção oposta do que sugere o parâmetro geral para esse grupo.

Figura 40 - Curva ajustada dos participantes do grupo homens sem conhecimento.

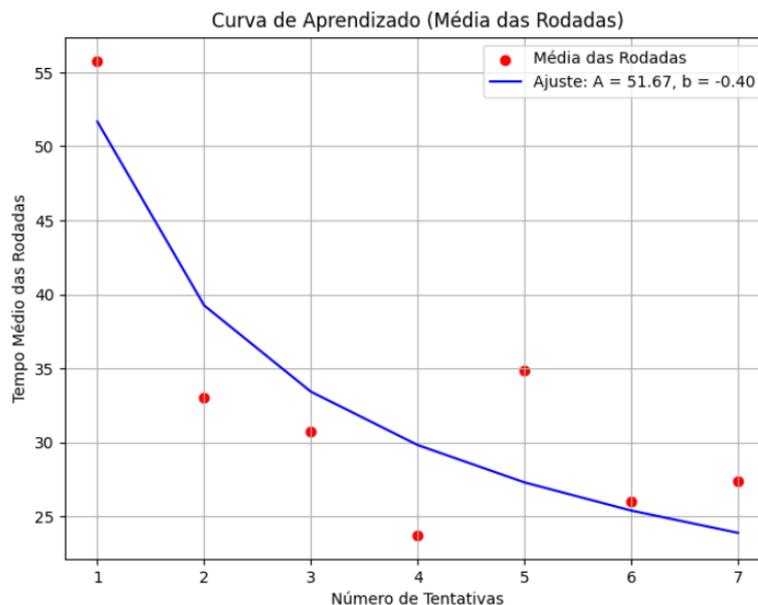


Fonte: O autor (2024).

4.1.4.4. Grupo mulheres sem conhecimento

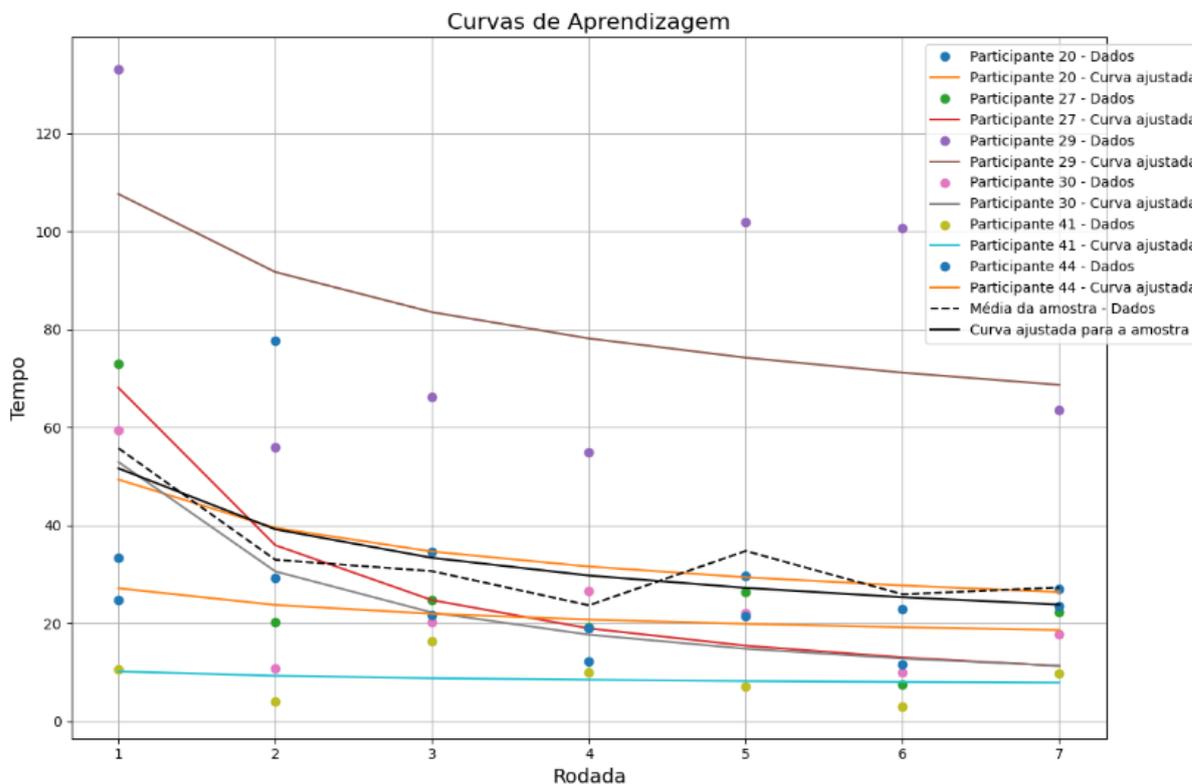
A Figura 41 mostra a curva de aprendizagem para o grupo mulheres sem conhecimento. A curva ajustada, apresentada na figura, se adapta bem aos dados do tempo médio de duração das rodadas para esse grupo. A curva apresenta parâmetros A e B, respectivamente iguais a 51,67 e -0,40, analisando a taxa de aprendizagem concluímos que o a melhoria ao longo de tempo acontece de forma gradual, porém não tão gradual e prolongada como a curva da amostra e nem não acentuada como a curva do grupo homens com conhecimento. Esse grupo apresenta menor tempo de execução da primeira e da última rodada, além de ter a mesma tendência de decréscimo apresentada pelos demais grupos.

Figura 41 - Curva ajustada dos participantes do grupo mulheres sem conhecimento.



A Figura 42 apresenta as curvas de aprendizado para todos os participantes pertencentes ao grupo mulheres sem conhecimento. Esse grupo apresenta pouca variabilidade, mas tem um participante que está muito acima dos outros, com tempo médio elevado em todas as rodadas. Esse grupo também apresenta uma baixa amplitude, se desconsiderar o participante acima da média, além disso o grupo tem os menores tempo iniciais e finais. As curvas de aprendizagem dos participantes desse grupo são suaves com pouco grau de inclinação, sugerindo pouca diferença entre as rodadas iniciais e finais, indicando que o platô de aprendizagem tem uma diferença relativamente pequena quando comparada com os tempos médios das rodadas iniciais. Assim como nos outros grupos, o grupo de mulheres sem conhecimento tem participantes com comportamento linear, o que indica pouco comprometimento com o experimento, uma vez que os participantes desse grupo não têm conhecimento prévio. Apesar do participante 29 ser o participante acima da média, ele possui uma das menores taxas de aprendizagem, e quando comparado ao participante 30, tem o dobro do tempo inicial e três vezes menos taxa de aprendizagem.

Figura 42 - Curva ajustada dos participantes do grupo mulheres sem conhecimento.



Fonte: O autor (2024).

4.1.4.5. Comparação entre grupos

Para comparar os quatro grupos tem se como base os dois parâmetros e o comportamento da curva, onde o parâmetro A representa o desempenho inicial e o parâmetro B a taxa de aprendizagem, como já comentando anteriormente. O grupo composto por homens com conhecimento, o parâmetro A desse grupo é 87,49, indicando o grupo com maior potencial de aprendizagem, no entanto, o parâmetro B desse grupo é o mais negativo de todos, sugerindo uma queda mais acentuada no desempenho ao longo do tempo. Comparando os dois parâmetros é possível identificar que os participantes desse grupo, embora comecem com um tempo de execução mais elevado aprendem mais rápidos, reduzindo o tempo de execução das rodadas seguintes de forma mais significativa.

O grupo de mulheres com conhecimento, apresenta parâmetro A menor que o grupo dos homens com conhecimento e parâmetro B menos negativo, sugerindo que apesar de ter um tempo inicial inferior ao grupo dos homens com conhecimento, o parâmetro B menos negativo sugere que a redução do desempenho acontece de forma mais gradativa que o no grupo homens com conhecimento, tendo maior estabilidade do que o grupo dos homens com conhecimento.

Os homens sem conhecimento têm parâmetro A igual a 69,29, que é superior ao das mulheres com conhecimento, mas inferior aos homens com conhecimento. Esse fato demonstra que mesmo se esse grupo possuir conhecimento sobre decisão multicritério o seu potencial de melhoria supera o grupo de mulheres com conhecimento. Em relação ao parâmetro B, que é igual a -0,45, é menos negativo que os grupos homens e mulheres com conhecimento, sugerindo uma redução menos acentuada.

Por fim, para o grupo das mulheres sem conhecimento foi verificado o menor valor para A e o maior para B. A curva desse grupo apresenta um comportamento mais estável quando comparado com os outros grupos.

Anteriormente foi sugerido duas hipóteses para a tendência de decrescimento, a primeira estava relacionada com a influência que uma rodada tem sobre a outra, já a segunda sugeria que o efeito do aprendizado causava a redução no tempo de execução. Ao estudar a taxa de aprendizado de cada grupo que compõem a amostra, foi visto que todas possuem um parâmetro B negativo, que indica uma redução causada pela taxa de aprendizagem. Dessa forma é possível sugerir que existe uma redução no tempo da avaliação intracritério causado pelo conhecimento.

Apesar de existir o feito da aprendizagem na redução dos tempos de cada rodada, e cada grupo ter sua taxa de aprendizagem, é preciso investigar se existe uma diferença estatisticamente significativa entre as medias dos participantes que compõem cada grupo. Para realizar essas comparações, será realizado um teste t para todas as combinações dos grupos.

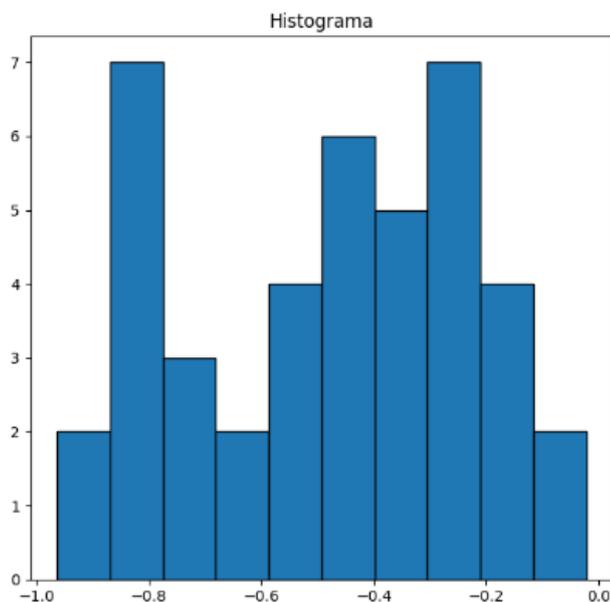
Antes de aplicar os testes t, é preciso testar os pressupostos de normalidade e homoscedasticidade, para os valores das taxas de aprendizado para cada participante. A Tabela 28, apresenta os testes de Shapiro-Wilk para normalidade e o de Levene para a homoscedasticidade, além disso a Figura 43 apresenta o histograma que está relacionado com a normalidade e a Figura 44 o gráfico dos resíduos que mostra a homogeneidade das variâncias. Analisando a tabela e comprovando com as figuras, constatamos que os dados obedecem às suposições de normalidade e homoscedasticidade, sendo possível aplicar o teste t.

Tabela 28 - Testes de normalidade e homoscedasticidade para a taxa de aprendizado.

Teste	p-valor	Interpretação
Shapiro-Wilk	0,116	Os dados são normalmente distribuídos,
Levene	0,758	As variâncias são homogêneas.

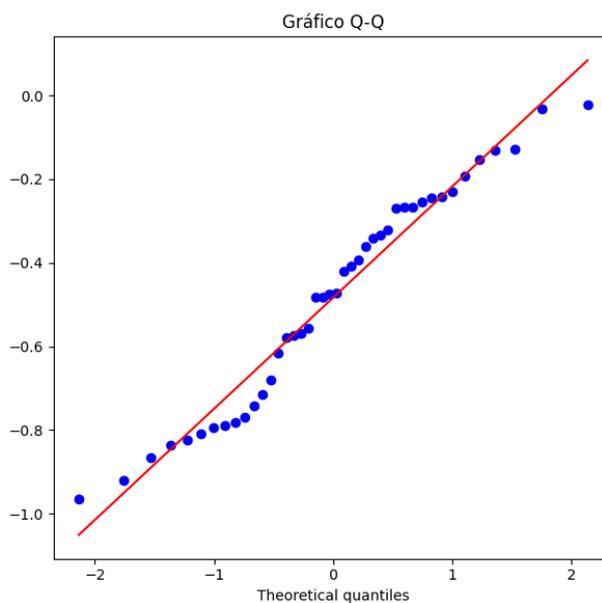
Fonte: O autor (2024).

Figura 43 - Histograma da taxa de aprendizagem.



Fonte: O autor (2024).

Figura 44 - Gráfico dos resíduos para a taxa de aprendizagem.



Fonte: O autor (2024).

O primeiro par de grupos que será testado é o grupo homens com conhecimento e homens sem conhecimento. A Tabela 29, apresenta o resultado do teste t para os grupos, o p-valor para esse teste foi de 0,072 e o nível de confiança adotado foi de 0,10, nesse caso como o p-valor é menor que o nível de significância rejeitamos a hipótese alternativa, concluindo que as médias entre os participantes desses dois grupos são diferentes.

Tabela 29 - Teste t para os grupos homens com conhecimento e homens sem conhecimento.

Homens com conhecimento x Homens sem conhecimento		teste-t para Igualdade de Médias						
		t	df	Sig.	Diferença média	Erro padrão	90% Intervalo de Confiança	
							Inferior	Superior
Aprendizagem	Variâncias iguais assumidas	-1,87	26	0,072	-0,17108	0,09115	-0,32655	-0,01562

Fonte: O autor (2024).

O segundo par de grupos a ser testado é o de mulheres com conhecimento e mulheres sem conhecimento. A Tabela 30, traz o resultado do teste t para a comparação desses dois grupos, o p-valor foi de 0,662 que é maior que o nível de significância indicando que devemos aceitar a hipótese nula de que não há diferença entre as médias desses dois grupos.

Tabela 30 - Teste t para os grupos mulheres com conhecimento e mulheres sem conhecimento.

Mulheres com conhecimento x Mulheres sem conhecimento		teste-t para Igualdade de Médias						
		t	df	Sig.	Diferença a média	Erro padrão	90% Intervalo de Confiança	
							Inferior	Superior
Aprendizagem	Variâncias iguais assumidas	-0,448	12	0,662	-0,07279	0,16258	-0,36256	0,21698

Fonte: O autor (2024).

O terceiro par de grupos a ser testado é o de Homens com conhecimento e mulheres com conhecimento. A Tabela 31, apresenta o resultado do teste t que compara os dois grupos, o p-valor associado a essa comparação é 0,524 que é maior que o nível de significância de 0,10. Assim aceitamos a hipótese nula de que a média dos participantes para esses grupos são iguais.

Tabela 31 - Teste t para os grupos homens com conhecimento e mulheres com conhecimento

Homens com conhecimento x Mulheres com conhecimento		teste-t para Igualdade de Médias						
		t	df	Sig.	Diferença média	Erro padrão	90% Intervalo de Confiança	
							Inferior	Superior
Aprendizagem	Variâncias iguais assumidas	-0,650	19	0,524	-0,07702	0,11850	-0,28192	0,12789

Fonte: O autor (2024).

O quarto par a ser testado são os grupos homens com conhecimento e mulheres sem conhecimento. A Tabela 32, apresenta o resultado do teste t, o valor de p para essa comparação foi de 0,099 que é menor que o nível de significância, assim tem se base para rejeitar a hipótese nula. Dessa maneira, concluímos que a diferença entre as médias desses grupos é estatisticamente significativa.

Tabela 32 - Teste t para os grupos homens com conhecimento e mulheres sem conhecimento.

Homens com conhecimento x Mulheres sem conhecimento		teste-t para Igualdade de Médias						
		t	df	Sig.	Diferença a média	Erro padrão	90% Intervalo de Confiança	
							Inferior	Superior
Aprendizagem	Variâncias iguais assumidas	-1,071	17	0,099	-0,14981	0,13991	-0,39320	0,09359

Fonte: O autor (2024).

O quinto par a ser testado são os grupos de mulher com conhecimento e homens sem conhecimento. A Tabela 33, apresenta o resultado do teste t entre os grupos, o valor p dessa comparação foi de 0,384 que é maior que o nível de significância, desse modo pode se aceitar a hipótese nula, de que as médias desses dois grupos são iguais.

Tabela 33 - Teste t para os grupos homens com conhecimento e mulheres sem conhecimento.

Mulheres com conhecimento x Homens sem conhecimento		teste-t para Igualdade de Médias						
		t	df	Sig.	Diferença média	Erro padrão	90% Intervalo de Confiança	
							Inferior	Superior
Aprendizagem	Variâncias iguais assumidas	0,889	21	0,384	0,09407	0,10577	-0,08793	0,27606

Fonte: O autor (2024).

Por fim, o sexto par a ser testado são os grupos de homens sem conhecimento e mulheres sem conhecimento. A Tabela 34, apresenta o resultado do teste t para os grupos, o valor de p para essa comparação foi de 0,867 que é muito maior que o valor do nível de significância

estabelecido. Dessa maneira, rejeitamos a hipótese nula e concluímos que as médias para esses grupos são iguais.

Tabela 34 - Teste para os grupos homens sem conhecimento e mulheres sem conhecimento.

Homens sem conhecimento x mulheres sem conhecimento		teste-t para Igualdade de Médias						
		t	df	Sig.	Diferença média	Erro padrão	90% Intervalo de Confiança	
							Inferior	Superior
Aprendizagem	Variâncias iguais assumidas	0,170	19	0,867	0,02128	0,12491	-0,19471	0,23727

Fonte: o Autor (2024).

Comparando os testes t e os parâmetros B de cada grupo, é possível perceber que apesar dos grupos terem suas respectivas taxas de aprendizagem, não houve uma diferença estatisticamente significativa entre a maioria das comparações. Avaliando as comparações entre os grupos homens com conhecimento (parâmetro B = -0,60) e homens sem conhecimento (parâmetro B = -0,45), houve uma diferença entre as médias estatisticamente significativa, além disso a diferença entre as taxas de aprendizagem desses grupos é de -0,15. A outra comparação que obteve uma diferença estatisticamente significativa foi entre os grupos homens com conhecimento e mulheres sem conhecimento (parâmetro B = -0,40), que tem uma diferença de -0,20 entre as taxas de aprendizagem. Todas as outras comparações obtiveram uma diferença nas taxas de aprendizagem inferiores a -0,15, o que de certa forma está alinhado com o resultado dos testes t que deram significativo e os que não deram significativo.

A média do parâmetro A, também foi comparada entre os grupos que compõem a amostra. Contudo, apenas constatou-se uma diferença estatisticamente significativa entre os grupos homens com conhecimento e mulheres sem conhecimento. A Tabela 35 traz o teste t que compara as médias do parâmetro A entre os grupos homens com conhecimento e mulheres sem conhecimento.

Tabela 35 – Teste t para os grupos homens com conhecimento e Mulheres sem conhecimento.

Homens com conhecimento x Mulheres sem conhecimento		teste-t para Igualdade de Médias						
		t	df	Sig.	Diferença média	Erro padrão	95% Intervalo de Confiança da Diferença	
							Inferior	Superior
Parâmetro A	Variâncias iguais assumidas	1,925	17	,071	35,79468	18,59820	-3,44410	75,0334

Fonte: O autor (2024).

4.1.5. Abordagem Random Forest

Como mostrado anteriormente, existe um efeito do conhecimento prévio sobre o aprendizado dos participantes ao longo do experimento. Apesar de haver uma diferença entre as taxas de aprendizagens dos grupos, é preciso entender como o gênero e o conhecimento se relacionam com a taxa de aprendizagem, contudo essa relação nem sempre é simples e direta, normalmente é uma relação complexa que os métodos tradicionais não conseguem captar.

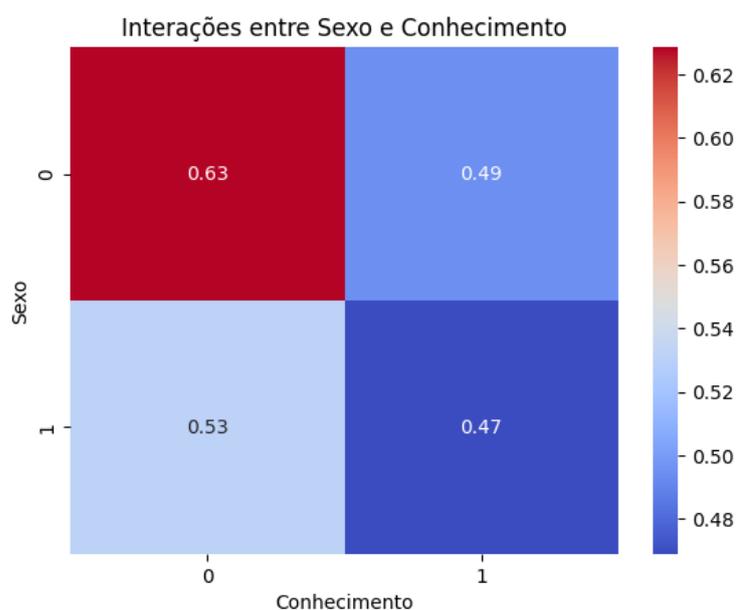
Tendo em vista a complexidade da relação entre as variáveis, esse trabalho se propõe a utilizar o modelo desenvolvido por Breiman (2001), denominado Random Forest (RF) e é uma extensão da abordagem tradicional das árvores decisão. O funcionamento desse modelo é dado por um conjunto de árvores de decisão, onde cada árvore utiliza uma amostra aleatória, do conjunto de dados inicial, para ser treinada. O processo de divisão do conjunto de dados inicial é chamado de bootstrap e tem como principal objetivo reduzir o número de overfitting. Partindo desse princípio Liaw e Wiener (2002), descrevem que para um output ser gerado o RF leva em consideração todos os resultados das árvores de decisão, utilizando a média das decisões ou a maioria as respostas.

Além de captar relações complexas entre as variáveis, outro aspecto que fez esse modelo ser escolhido é sua não sensibilidade a outliers, como sugerido por Hastie (2009), onde é afirmado que o random forest tem uma grande flexibilidade que permite sua aplicação em diversas problemáticas sem perder a robustez, isso é dado pela sua não sensibilidade a dados faltantes, presença e outliers.

A Figura 45, apresenta a matriz de interações, essa matriz expressa a interação entre as variáveis sexo, em que 0 representa o sexo feminino e 1 o sexo masculino, e a variável conhecimento, em que 0 representa não ter conhecimento e 1 representa ter conhecimento. A matriz avalia a força da interação entre as variáveis para descrever a variável dependente, que

nesse caso é a taxa de aprendizagem. Analisando a matriz, vemos que o modelo considerou a interação entre a variável sexo com valor 0 e a variável conhecimento com valor 0 a mais forte, por outro lado a interação entre sexo com valor 1 e conhecimento com valor 0, foi considerada a mais fraca. De maneira geral, a força das interações é traduzida na influência que essa combinação causa nas respostas certas do modelo.

Figura 45 - Matriz de interação entre gênero e conhecimento.



Fonte: O autor (2024).

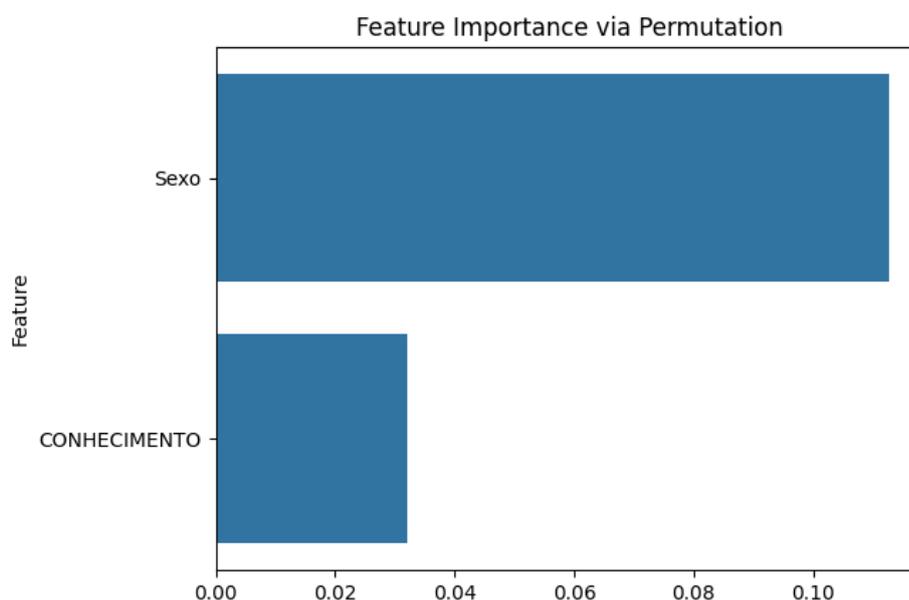
Apesar da matriz de interação apresentar a importância da combinação das classes de cada variável, ela não indica a importância que a variável sexo e conhecimento tem sobre o modelo. A Tabela 36 apresenta a porcentagem que cada variável tem nos resultados do RF, é possível ver que ambas as variáveis se mostram importantes, porém o sexo tem uma contribuição um pouco maior que o conhecimento. Para confirmar essa indicação é feito um teste de permutação entre as variáveis, esse teste avalia o a influência que a permuta de cada uma das variáveis tem sobre o modelo. A Figura 46 apresenta o quanto o modelo sofre quando ocorre a permutação das variáveis sexo e conhecimento, analisando a figura, vemos que o modelo tem uma perda de aproximadamente 0,10 na sua performance quando a variável sexo é permutada, enquanto a permutação da variável conhecimento afeta apenas em aproximadamente 0,03.

Tabela 36 - Importância das variáveis para o modelo

Variável	Contribuição
Sexo	51,39%
Conhecimento	48,61%

Fonte: O autor (2024).

Figura 46 - Impacto da permutação das variáveis.



Fonte: O autor (2024).

A importância das variáveis sexo e conhecimento, assim como a interação de cada resultado, já foram descritas anteriormente, mas é possível aprofundar um pouco mais e tentar entender a influência que cada sexo e cada nível de conhecimento tem nas previsões do modelo. Para isso é preciso decompor as variáveis sexo e conhecimento em 4 variáveis binárias, essa técnica é conhecida como one-hot encoding, essa técnica é explicada por Bishop (2006) como sendo a representação numérica das variáveis categóricas de modo que os algoritmos possam processar com mais assertividade. O Quadro 4 apresenta a como será feita a decomposição das variáveis e a descrição das novas variáveis criadas.

Quadro 4 - Decomposição das variáveis.

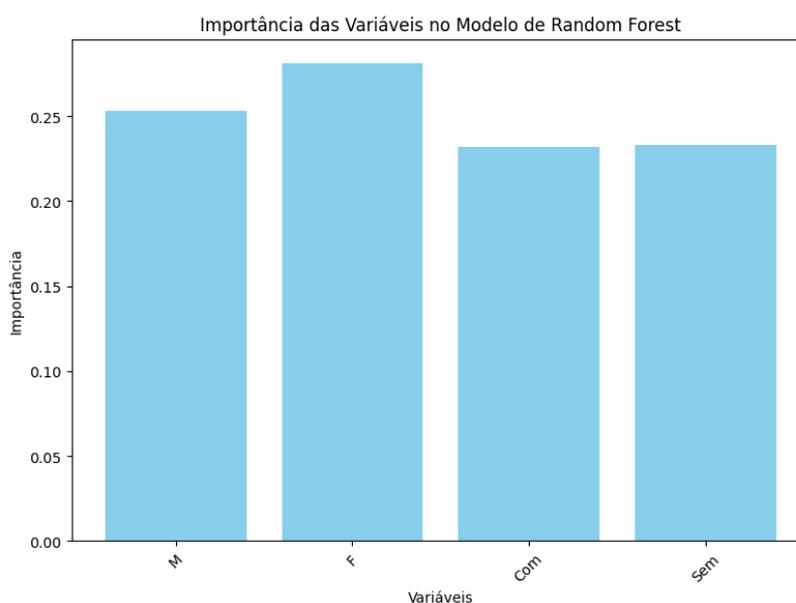
Variável original	Variável Binária	Descrição
Sexo	Sexo feminino	Receberá 1 caso o participante for do sexo feminino e 0 se não for.

	Sexo Masculino	Receberá 1 se o participante for do sexo masculino e 0 se não for.
Conhecimento	Sem conhecimento	Receberá 1 caso o participante não tenha conhecimento e 0 se tiver.
	Com conhecimento	Receberá 1 caso o participante tenha conhecimento e 0 caso não tenha.

Fonte: O autor (2024).

A Figura 47 apresenta a importância das variáveis decompostas em porcentagem, contudo a Tabela 37 apresenta os valores exatos de cada porcentagem. Analisando o gráfico e a tabela, concluímos que a diferença entre as variáveis é pequena, porém o sexo feminino ajuda o modelo a descrever melhor os resultados, já em relação ao conhecimento a diferença é tão pequena que podemos considerar que as duas variáveis têm igual importância.

Figura 47 - Importância das variáveis decompostas



Fonte: O autor (2024).

Tabela 37 - Porcentagem do impacto de cada variável decomposta

Variável	Importância
Sexo masculino	25,34%
Sexo feminino	28,12%
Com conhecimento	23,22%
Sem conhecimento	23,32%

Fonte: O autor (2024).

4.2. Tempo de fixação

4.2.1. Análise descritiva

Utilizando o eye-tracking, foi possível coletar os dados sobre a quantidade de fixações que os participantes tiveram, em cada uma das áreas de interesse, previamente definidas, ao longo do experimento. Assim, sendo necessário a análise descritiva do tempo de fixação de cada grupo nas áreas de interesse previamente estabelecidas, a fim de identificar padrões de comportamento.

Analisando a Tabela 38, é possível sugerir que a área Choice obteve a maior média de fixação (19,58) e apresentou uma variação considerável (Desvio Padrão de 5,89), indicando que os participantes com maior conhecimento dedicaram uma atenção significativa, mas de forma variável. A área Graph teve uma média menor (7,99) e um desvio padrão alto (5,04), sugerindo que o engajamento visual foi inconsistente, possivelmente devido à complexidade ou relevância variável do conteúdo. A área InstB apresentou a menor média de fixação (0,45) e uma baixa variação, o que indica que não gerou muita atenção, reforçando sua menor relevância. Por outro lado, InstD teve uma média moderada (9,54), mas com uma grande dispersão (Desvio Padrão de 7,01), sugerindo que alguns participantes prestaram mais atenção a essa área do que outros. A grande variação observada em Choice e InstD sugere que há uma diferença na distribuição da atenção, possivelmente influenciadas por aspectos cognitivos ou preferências pessoais. Esses padrões sugerem que a região Choice despertou maior interesse, enquanto Graph e InstD apresentaram uma atenção mais dispersa, e InstB foi a menos atrativa para esse grupo.

Tabela 38 - Tempo médio de fixação do grupo homens com conhecimento.

	Choice	Graph	InstB	InstD
Rodada 1	31,36	18,58	1,38	23,28
Rodada 2	16,48	6,01	0,53	9,51
Rodada 3	18,30	7,44	0,16	13,38
Rodada 4	18,92	9,75	0,37	7,21
Rodada 5	20,17	5,28	0,37	6,91
Rodada 6	12,03	3,94	0,21	3,59
Rodada 7	19,83	4,96	0,16	2,92

Fonte: O autor (2024).

Já a Tabela 39, apresenta os dados do grupo mulheres com conhecimento, os dados mostram que a área Choice recebeu a maior média de fixação (17,35) e teve uma variação

considerável (Desvio Padrão de 6,29), o que sugere que foi a região que mais chamou a atenção, embora de forma um pouco instável entre as participantes. Já a área Graph teve uma fixação mais baixa (3,90) e pouca variação, indicando que o envolvimento visual foi mais uniforme, mas talvez com uma relevância menor para o grupo. InstB teve a média mais baixa (0,73) e uma variação relativamente alta (Desvio Padrão de 1,04), sugerindo que, enquanto alguns participantes prestaram atenção nessa área, outras quase não a notaram. Por fim, InstD teve uma média moderada (9,72), mas a maior variação (Desvio Padrão de 8,20), o que indica que houve uma diferença significativa na forma como cada participante focou nessa região. O padrão geral aponta para uma forte concentração de atenção em Choice, uma atenção mais variável em InstD e um engajamento mais baixo com Graph e InstB, sugerindo que as participantes estavam mais interessadas nas informações centrais.

Tabela 39 - Tempo médio de fixação do grupo mulheres com conhecimento.

	Choice	Graph	InstB	InstD
Rodada 1	28,56	7,09	3,05	25,53
Rodada 2	19,30	3,97	0,32	14,84
Rodada 3	13,62	4,57	0,46	10,24
Rodada 4	15,43	3,45	0,21	6,35
Rodada 5	14,91	4,23	0,64	5,84
Rodada 6	8,83	1,59	0,24	2,65
Rodada 7	20,81	2,40	0,19	2,59

Fonte: O autor (2024).

A Tabela 40 apresenta as médias de fixação do grupo homens sem conhecimento, e indica que a área Choice foi a mais observada, com uma média de fixação de 16,75, mas com uma variação considerável (Desvio Padrão de 5,54), sugerindo que foi um ponto de destaque, embora com diferenças no comportamento de atenção entre os participantes. A área Graph, com média de 4,12 e uma variação menor (Desvio Padrão de 1,31), parece ter atraído uma atenção mais uniforme, mas de forma menos relevante para o grupo. Já InstB teve a menor média (0,42) e uma baixa amplitude (0,97), indicando que essa área gerou pouquíssimo interesse. Por outro lado, InstD, com uma média de 12,14 e uma grande variação (Desvio Padrão de 7,44), mostra que houve uma atenção bem diferente entre os participantes, com alguns focando mais nessa região do que outros. A ampla variação em InstD (22,13) e Choice (17,39) destaca as diferenças individuais na distribuição da atenção. No geral, o padrão sugere que a atenção foi mais concentrada em Choice, com um foco moderado e variável em InstD, enquanto Graph e InstB receberam pouca atenção.

Tabela 40 - Tempo médio de fixação do grupo homens sem conhecimento.

	Choice	Graph	InstB	InstD
Rodada 1	18,66	6,37	1,17	26,38
Rodada 2	11,80	3,59	0,42	14,74
Rodada 3	16,24	4,34	0,23	10,71
Rodada 4	18,58	4,93	0,24	13,62
Rodada 5	16,86	3,87	0,22	10,44
Rodada 6	8,85	2,13	0,20	4,82
Rodada 7	26,24	3,60	0,46	4,25

Fonte: O autor (2024).

Observando a Tabela 41, dos dados do grupo mulheres sem conhecimento, tem-se que a área Choice teve uma média de fixação de 10,64, com uma grande variação (Desvio Padrão de 3,57), o que indica que, embora tenha sido um ponto de atenção, o nível de fixação foi inconsistente ao longo das sessões. A área Graph, com uma média de 4,18 e desvio padrão de 2,07, também despertou um interesse moderado, mas com uma variação significativa, sugerindo que a atenção voltada para essa região foi irregular. InstB, por sua vez, teve a menor média (0,23) e um desvio padrão muito baixo (0,30), o que mostra que essa área foi praticamente ignorada pela maioria das participantes. Já InstD, com média de 10,62, teve a maior dispersão (Desvio Padrão de 3,83), indicando que alguns participantes prestaram muita atenção nessa área, enquanto outras praticamente a desconsideraram, o que resultou em grande variação no tempo de fixação. A ampla variação observada em Choice (9,49) e InstD (10,19) sugere diferenças individuais na forma como essas áreas foram percebidas e no quanto foram consideradas relevantes. Em resumo, a atenção se concentrou principalmente em Choice e InstD, enquanto Graph e InstB atraíram bem menos atenção.

Tabela 41 - Tempo médio de fixação do grupo mulheres sem conhecimento.

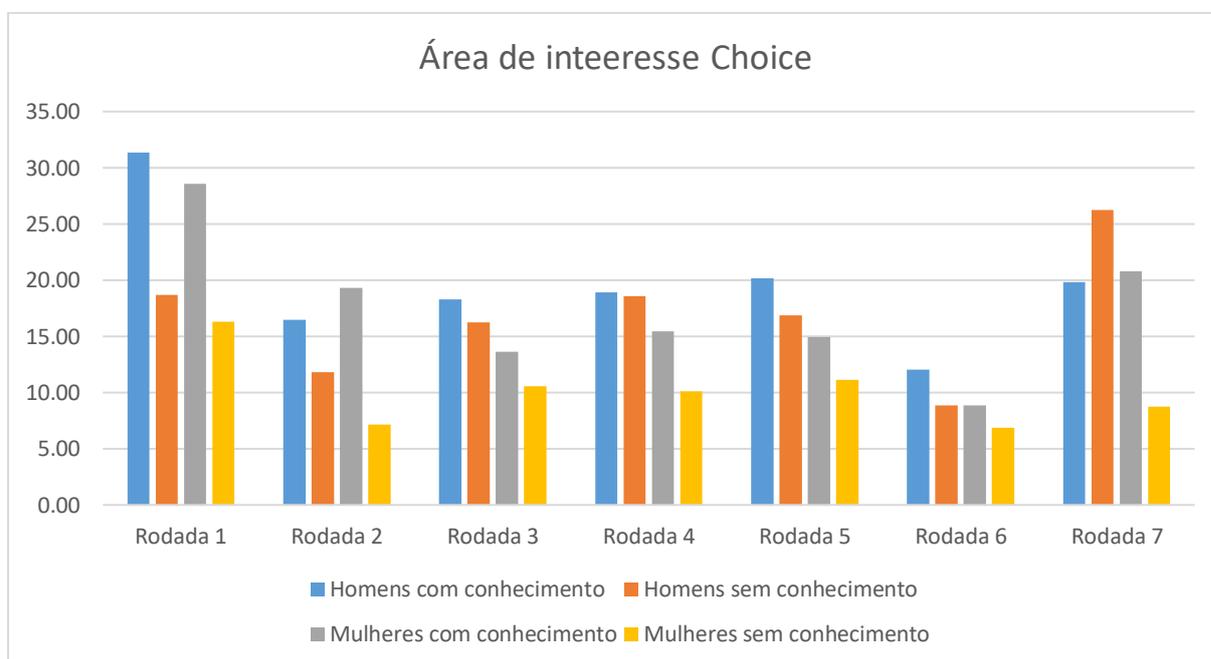
	Choice	Graph	InstB	InstD
Rodada 1	16,32	9,38	0,96	16,68
Rodada 2	7,17	3,61	0,15	9,98
Rodada 3	10,59	4,05	0,23	8,75
Rodada 4	10,08	3,13	0,01	6,18
Rodada 5	11,11	2,61	0,25	16,41
Rodada 6	6,89	2,72	0,09	12,79
Rodada 7	8,76	3,64	0,09	8,65

Fonte: O autor (2024).

Para todos os grupos as médias do tempo de fixação na área de interesse InstD diminuí ao longo das rodadas, com exceção do grupo mulheres sem conhecimento. Para esse grupo o tempo de fixação nessa área de interesse diminuí nas primeiras rodadas, mas torna a aumentar nas rodadas finais. Por outro lado, a área de interesse Choice foi a que teve, de forma geral, a maior média de tempo de fixação. No entanto, nos grupos com conhecimento, a atenção nessa área é a maior independente da rodada observada, enquanto nos grupos sem conhecimento a fixação na área Choice é menor que a área InstD, nas rodadas iniciais. Já a área InstB é a que tem o menor tempo de fixação, e possui uma tendência de decrescimento em todos os grupos.

A Figura 48, aborda uma análise comparativa entre os grupos na área de interesse Choice. Com exceção da rodada 2, todas as outras rodadas tiveram maior tempo de fixação por participantes do sexo masculino, além disso os participantes do sexo masculino que possuam conhecimento tiveram o maior número de fixação em cinco das sete rodadas. Por outro lado, o grupo de mulheres sem conhecimento, teve o menor tempo de fixação em todas as rodadas.

Figura 48 - Comparação da fixação na área Choice.

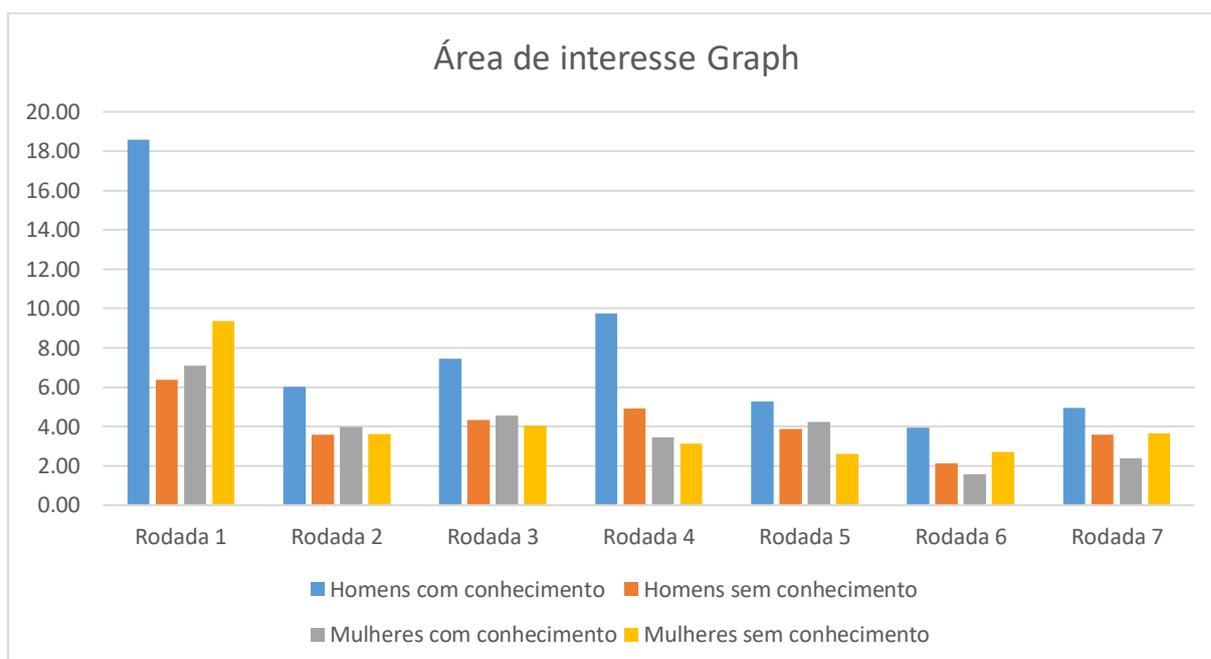


Fonte: O autor (2024).

De forma análoga, a Figura 49 apresenta a comparação, entre os grupos, do tempo médio de fixação. O grupo homens com conhecimento tiveram a maior fixação média em todas as rodadas, os grupos de mulheres sem conhecimento e mulheres com conhecimento obtiveram o

segundo maior tempo médio de fixação em três rodadas. Por outro lado, o grupo homens sem conhecimento teve o segundo maior tempo médio de fixação em apenas uma rodada, nas demais ele oscilou entre a terceira e quarta posição.

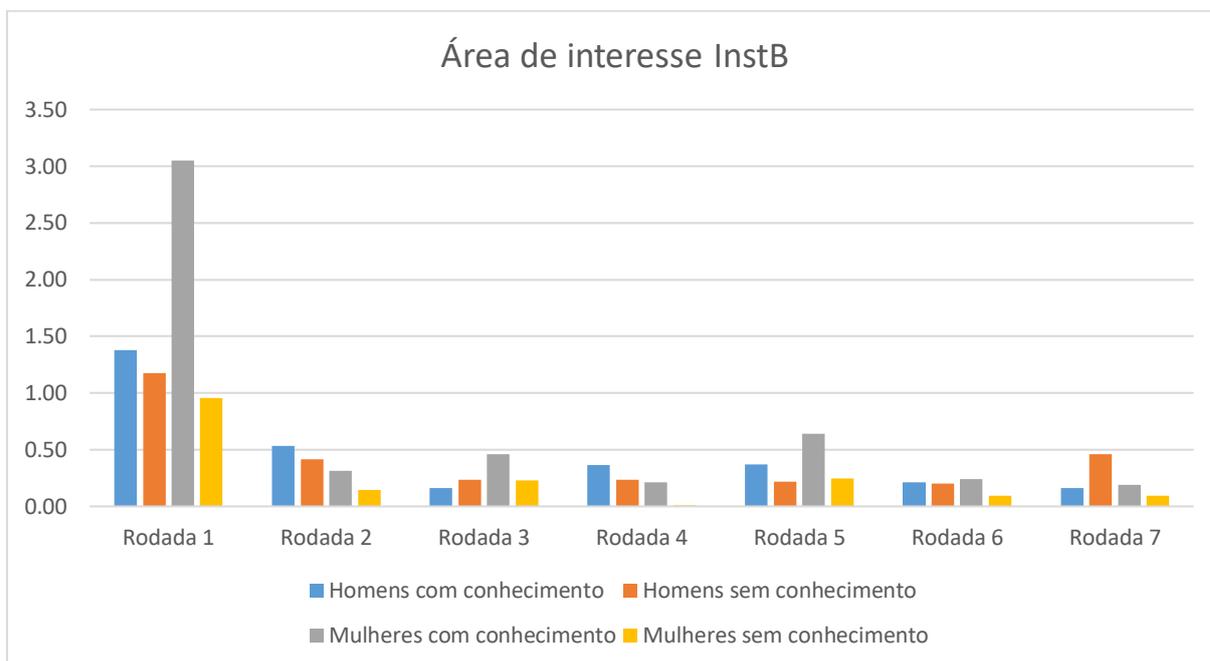
Figura 49 - Comparação da fixação na área Graph.



Fonte: O autor (2024).

De forma análoga, a Figura 50 compara os grupos para a área de interesse InstB. O grupo mulheres com conhecimento foi o grupo que mais fixou a atenção nessa área de interesse, já os homens com conhecimento ocupam a segunda posição em cinco das sete rodadas. O grupo de homens sem conhecimento ocupou a primeira posição apenas na rodada sete e a segunda posição nas rodadas dois e três. O grupo de mulheres sem conhecimento teve o menor tempo médio de fixação em cinco das sete rodadas e ocupou a terceira posição em duas das sete rodadas.

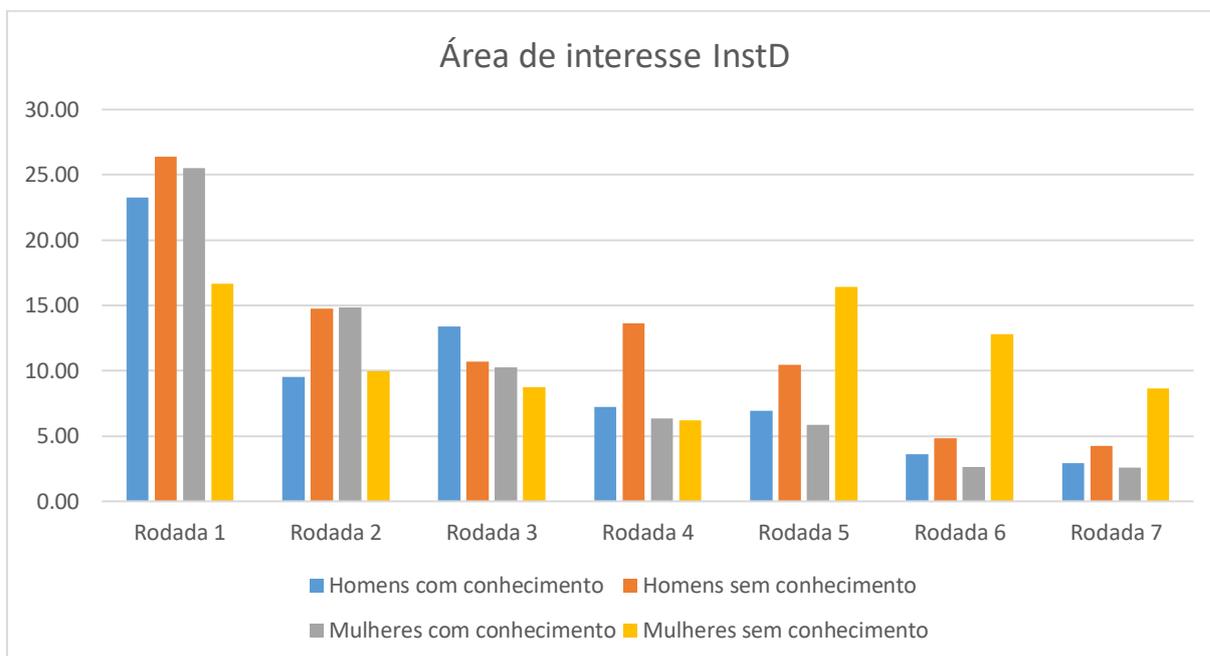
Figura 50 - Comparação da fixação na área InstB.



Fonte: O autor (2024).

Por fim, a Figura 54 compara os tempos médios de fixação na área de interesse InstD. O grupo de mulheres sem conhecimento ocupou a primeira posição nas rodadas finais, por outro lado o grupo de mulheres com conhecimento tiveram o terceiro maior tempo de fixação em cinco das sete rodadas e a primeira posição na rodada 2. O grupo de homens sem conhecimento teve o maior tempo de fixação nessa área em três rodadas e ocupou a segunda posição nas demais. O grupo homens com conhecimento, por sua vez, ocupou a primeira posição na rodada terceira rodada e oscilou entre a segunda e quarta posição nas demais.

Figura 51 - Comparação da fixação na área InstD.



Fonte: O autor (2024).

De maneira geral, os grupos de homens com e sem conhecimento tendem a olhar mais para as instruções nas primeiras rodadas, porém esse tempo diminui com o avanço do experimento ao mesmo passo que o tempo de fixação aumenta nas áreas de interesse das escolhas e do gráfico. Já os grupos de mulheres com e sem conhecimento distribuem melhor sua atenção nas áreas de interesse, contudo focando mais na área das alternativas e nas instruções D. Os grupos sem conhecimento tendem a ignorar a área de interesse B ao longo do experimento, uma vez que seu tempo de fixação diminui de forma acentuada até se aproximar de zero em algumas rodadas. Por outro lado, os grupos com conhecimento diminuem a fixação nessa área, mas ainda mante um tempo de atenção considerável, apesar de ser pouco.

Utilizando o método da análise de variância (ANOVA), um modelo foi criado para tentar explicar a relação e a influência que o gênero do participante e o fato dele ter ou não conhecimento acerca de decisão multicritério, pode influenciar se o participante tem mais propensão a olhar para uma das áreas de interesse previamente definidas.

4.2.2. Análise de variância para o tempo de fixação

Devido as diferenças entre os tempos de fixação, apresentados na sessão anterior, se faz necessário entender se existe uma influência do grau de conhecimento e do gênero. Para verificar essa condição foi aplicado uma ANOVA para cada área de interesse em cada rodada,

verificando se existe uma influência, do gênero e do grau de conhecimento, estatisticamente significativa que explique a fixação nas áreas de interesse.

Como já discutido anteriormente, a aplicação da ANOVA exige que dois pressupostos sejam atendidos (normalidade dos dados e homoscedasticidade). Os dados da duração da fixação em cada rodada foram submetidos aos testes de Shapiro-Wilk para normalidade e o teste de Levene para homoscedasticidade, todas as rodadas atenderam aos pressupostos. Assim, os próximos subtópicos dessa sessão apresenta os resultados significativos da ANOVA para cada rodada e área de interesse.

4.2.2.1. Rodada 1

Na primeira rodada identificou-se uma diferença estatisticamente significativa entre as médias dos grupos na área Choice e Graph. A Figura 52 apresenta o resultado da ANOVA para o tempo de fixação na área Choice. Analisando essa figura, constatou-se que existe uma diferença entre as médias dos grupos causada pela variável conhecimento. Através da comparação das médias, nota-se que ter conhecimento impacta mais no tempo de fixação do que não ter conhecimento. Já a Figura 53 traz a ANOVA para área de interesse Graph, dessa maneira é possível notar que a combinação entre o gênero e o conhecimento causou uma diferença significativa no tempo de fixação dessa área. Assim, ser do gênero masculino e ter conhecimento faz com que o participante mantenha sua atenção por mais tempo nessa área.

Figura 52 - ANOVA da área Choice na primeira rodada.

Origem	Tipo III Soma dos Quadrados	df	Quadrado Médio	Z	Sig.
Modelo corrigido	1651,705 ^a	3	550,568	2,187	,105
Interceptação	20692,619	1	20692,619	82,186	,000
Genero	60,595	1	60,595	,241	,627
Conhecimento	1429,557	1	1429,557	5,678	,022
Genero * Conhecimento	,480	1	,480	,002	,965
Erro	9567,541	38	251,777		
Total	35700,103	42			
Total corrigido	11219,246	41			

Fonte: O autor (2024).

Figura 53 - ANOVA da área Graph na primeira rodada.

Origem	Tipo III Soma dos Quadrados	df	Quadrado Médio	Z	Sig.
Modelo corrigido	1204,235 ^a	3	401,412	3,877	,016
Interceptação	3940,349	1	3940,349	38,053	,000
Genero	165,343	1	165,343	1,597	,214
Conhecimento	226,604	1	226,604	2,188	,147
Genero * Conhecimento	483,136	1	483,136	4,666	,037
Erro	3934,831	38	103,548		
Total	9961,137	42			
Total corrigido	5139,066	41			

Fonte: O autor (2024).

4.2.2.2. Rodada 2

Na segunda rodada, houve uma diferença significativa apenas na área Choice, que se destaca através da Figura 54, que aborda o teste da ANOVA. Nessa rodada a variável conhecimento causou uma diferença estatisticamente significativa no tempo de fixação dos grupos nessa área. A análise conjunta entre a ANOVA e as médias de cada grupo, destacou que ter conhecimento aumenta a fixação do participante nessa área de interesse.

Figura 54 - ANOVA da área Choice na segunda rodada.

Origem	Tipo III Soma dos Quadrados	df	Quadrado Médio	Z	Sig.
Modelo corrigido	657,813 ^a	3	219,271	1,794	,165
Interceptação	6885,804	1	6885,804	56,347	,000
Genero	7,553	1	7,553	,062	,805
Conhecimento	649,863	1	649,863	5,318	,027
Genero * Conhecimento	127,659	1	127,659	1,045	,313
Erro	4643,708	38	122,203		
Total	13551,170	42			
Total corrigido	5301,520	41			

Fonte: O autor (2024).

4.2.2.3. Rodada 3

Na terceira rodada, a ANOVA apresentada na Figura 55, mostrou que a variável gênero tem um impacto estatisticamente significativo na média do tempo de fixação dos grupos. Ao analisar as médias, notou-se que ser do gênero masculino leva o participante a passar mais tempo olhando para essa área de interesse.

Figura 55 - ANOVA da área Choice na terceira rodada.

Origem	Tipo III Soma dos Quadrados	df	Quadrado Médio	Z	Sig.
Modelo corrigido	283,406 ^a	3	94,469	1,311	,285
Interceptação	7930,703	1	7930,703	110,060	,000
Genero	245,533	1	245,533	3,407	,073
Conhecimento	59,792	1	59,792	,830	,368
Genero * Conhecimento	2,178	1	2,178	,030	,863
Erro	2738,206	38	72,058		
Total	13207,194	42			
Total corrigido	3021,611	41			

Fonte: O autor (2024).

4.2.2.4. Rodada 4

Na quarta rodada, também ocorre o uma diferença entre as médias causada pela variável gênero, como pode ser visto na Figura 56. A análise conjunta entre a média dos grupos e a ANOVA, indicou que o gênero masculino tem efeito maior sobre a fixação nessa área que o gênero feminino.

Figura 56 – ANOVA da área Choice na quarta rodada.

Origem	Tipo III Soma dos Quadrados	df	Quadrado Médio	Z	Sig.
Modelo corrigido	391,692 ^a	3	130,564	2,154	,109
Interceptação	9121,248	1	9121,248	150,511	,000
Genero	330,440	1	330,440	5,453	,025
Conhecimento	74,225	1	74,225	1,225	,275
Genero * Conhecimento	57,558	1	57,558	,950	,336
Erro	2302,872	38	60,602		
Total	14648,647	42			
Total corrigido	2694,565	41			

Fonte: O autor (2024).

4.2.2.5. Rodada 5

Não foi observado nenhuma diferença estatisticamente significativa entre as médias dos grupos, nas áreas de interesse, nessa rodada.

4.2.2.6. Rodada 6

Não foi observado nenhuma diferença estatisticamente significativa entre as médias dos grupos, nas áreas de interesse, nessa rodada.

4.2.2.7. Rodada 7

A Figura 57, mostra que na sétima rodada ocorre uma diferença estatisticamente significativa entre as médias dos grupos, causado pela variável gênero e pela combinação entre a variável gênero e conhecimento. Dessa maneira, indicando que o tempo de fixação é influenciado pelo gênero masculino, como também é influenciada por ser do gênero masculino e ter conhecimento.

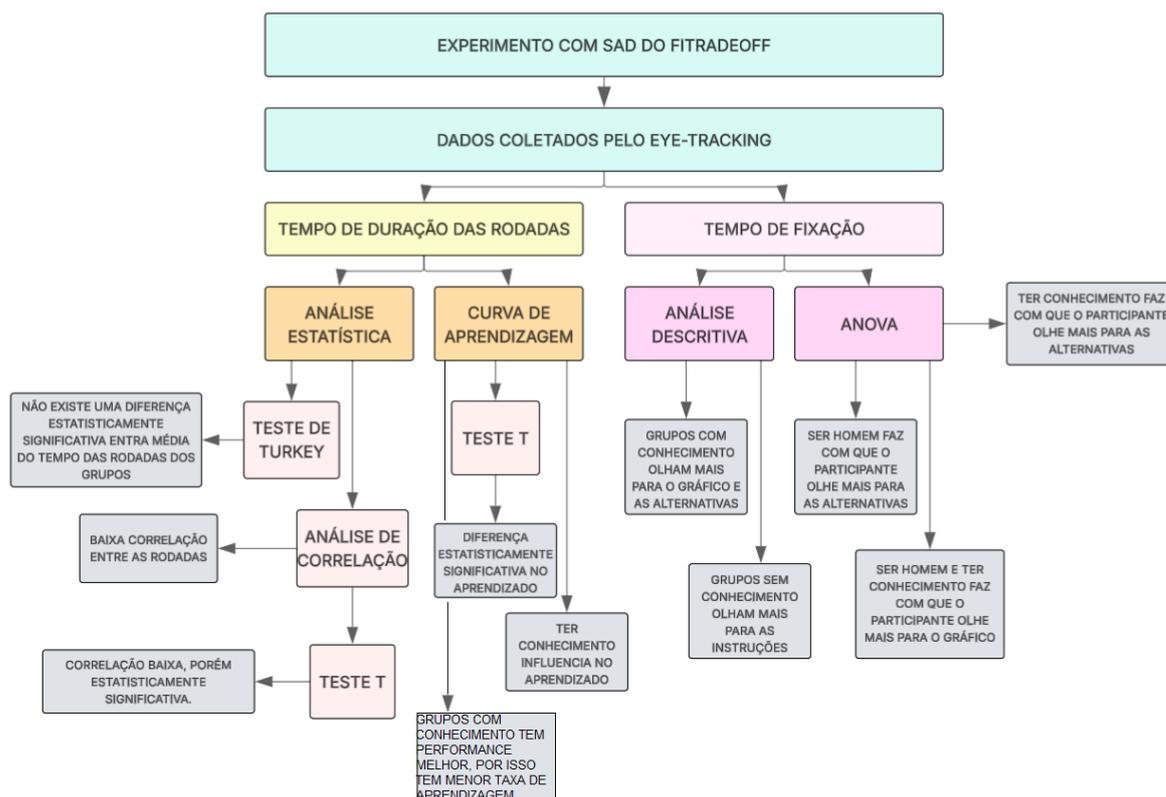
Figura 57 - ANOVA da área Choice na setima rodada.

Origem	Tipo III Soma dos Quadrados	df	Quadrado Médio	Z	Sig.
Modelo corrigido	1040,420 ^a	3	346,807	2,389	,084
Interceptação	12440,549	1	12440,549	85,701	,000
Genero	486,699	1	486,699	3,353	,075
Conhecimento	135,989	1	135,989	,937	,339
Genero * Conhecimento	626,301	1	626,301	4,315	,045
Erro	5516,136	38	145,161		
Total	23360,556	42			
Total corrigido	6556,556	41			

Fonte: O autor (2024).

Devido a extensa discussão que foi realizada nessa dissertação se faz necessário compilar todos os resultados, implicações e *insights*, de modo que seja simples identificar e compreender as informações. Assim, a Figura 58 apresenta todos os resultados e implicações obtidos ao longo da elaboração desse trabalho.

Figura 58 - Resumo dos resultados obtidos nesse trabalho



Fonte: O autor (2024)

5. CONCLUSÃO E SUGESTÕES PARA TRABALHOS FUTUROS

5.1. Conclusões

Os dados coletados através do experimento, ajudou a sugerir padrões de comportamento associados ao gênero e ao grau de conhecimento, na etapa de avaliação intracritério com SAD do FITradeoff. Permitindo avaliar especificamente o aprendizado por meio da curva de aprendizagem e da alocação da atenção pelos dados da duração da fixação nas áreas de interesse previamente definidas.

Através da curva de aprendizagem, de maneira geral, observou-se um percentual de aprendizado de aproximadamente 70%, indicando uma boa taxa de aprendizado ao longo do processo, com tempo inicial de 74,03 segundos que é considerado relativamente moderado.

Olhando de forma mais específica para os quatro grupos que compõem a amostra, ficou evidente que o conhecimento foi uma variável relevante para a redução da taxa de aprendizagem. Ter conhecimento prévio, reduziu o percentual do aprendizado de 73% para 66% no caso dos homens e de 75% para 69% no caso das mulheres. De maneira geral, os participantes do gênero masculino tiveram uma performance melhor que os do gênero feminino, apesar dessa diferença não ser tão acentuada. Por outro lado, o tempo gasto inicialmente foi mais elevado nos participantes do gênero masculino do que no gênero feminino. Além disso o conhecimento também foi um fator importante para o aumento no tempo de execução da primeira rodada, quando comparado com o grupo sem conhecimento.

A análise da curva de aprendizagem também sugere um padrão onde, o gênero masculino tende a aprender mais rápido, porém gastam mais tempo nas rodadas iniciais. Já o gênero feminino tende a ser mais constantes ao longo do processo, gastando menos tempo nas rodadas iniciais em comparação as últimas, mantendo o tempo mais uniforme quando comparado ao gênero masculino. O conhecimento, por sua vez, influencia no aprendizado na execução do experimento, reduzindo o tempo necessário para realizar as etapas decaio com o desenrolar do experimento.

Esses resultados são comprovados pelos testes t, que indicaram uma diferença estatisticamente significativa entre o tempo de execução dos grupos. As análises sugeriram uma diferença na média do tempo de realização da avaliação intracritério entre os grupos, homens com conhecimento e homens sem conhecimento e entre os grupos homens com conhecimento e mulheres sem conhecimento. Apoiando, dessa maneira, os resultados das análises descritivas.

Quanto ao estudo da atenção com os dados da duração da fixação, as áreas de interesse Choice e InstD, que tratam respectivamente das alternativas e das instruções, foram as áreas mais fixadas pelos grupos, seguidas pelas áreas Graph e InstB. Esses resultados indicam que a instrução do processo e uma área foca (Choice e Graph) no processo elicitatório são as mais buscadas pelos participantes. Para os grupos de homens com conhecimento e mulheres com conhecimento a área choice tende a ser a mais visualizada do começo ao fim do processo. Já nos outros acontece uma troca, nas primeiras rodadas, as áreas de instruções são mais procuradas e nas rodadas finais a área choice passar a ter mais foco de atenção. Esses fatos corroboram para sugerir que o conhecimento muda a perspectiva da atenção desde o primeiro momento do processo. Garantindo que as áreas que são mais importantes para a elicitación (Choice e Graph) sejam mais fixadas. Sugerindo, dessa maneira, que os grupos sem conhecimento levam metade do processo para aprender a fixar sua atenção para as áreas mais importante do processo elicitatório.

Ao comparar cada área de interesse ao longo do experimento, os resultados sugeriram que existe uma influência do gênero sobre a atenção nas áreas de interesse Choice. As análises indicaram que o gênero masculino tende a fixar mais a atenção na área Choice do que o gênero feminino. Em alguns casos o grau de conhecimento apoia o gênero na duração do tempo de fixação na área Choice, e na área Graph. Esses resultados foram confirmados pela ANOVA para o gênero na área Choice nas rodadas 1, 2, 3, 4 e 7, e para a área Graph na rodada 2. De maneira geral, os resultados então sugerem, que o gênero masculino tende a ter um perfil mais analítico e direto, enquanto o gênero feminino tem um perfil mais holístico e conservador.

Os experimentos comportamentais contribuem fortemente para o aperfeiçoamento dos SADs, o FITradeoff utiliza ferramentas da neurociência, como eletroencefalograma, eye-tracking, monitoramento cardíaco e a resposta galvânica da pele, para aperfeiçoar não só sua interface, mas também a experiência dos usuários e os algoritmos do método, para entregar resultados mais alinhados com as preferências do decisor. Este trabalho traz a utilização do eye-tracking para captação de dados, assim como meio de avaliar a interface do SAD.

5.2. Insights e Sugestões de Melhoria

Os resultados desse estudo são sugestivos em vários aspectos relacionados ao grau de conhecimento e ao gênero. As implicações desse estudo fornecem insights valiosos para os desenvolvedores de sistemas, em especial os desenvolvedores de SAD do FTtradeoff. Outra contribuição dessa obra, é voltada para os analistas de decisão responsáveis pelo suporte metodológico durante o processo decisório. É importante entender como o gênero e o grau de conhecimento impactam o processo elicitatório, a fim de oferecer suporte mais direcionado e efetivo para esses atores.

Com base no perfil analítico e direto do gênero masculino e no perfil holístico e conservador do gênero feminino, pode-se sugerir o suporte para o gênero masculino seja mais focada em aspectos números e direcionais, tomando proveito da característica analítica desse grupo. Já para o gênero feminino, sugere-se que o suporte seja mais voltado para uma contextualização os resultados e as informações apresentadas. Para os desenvolvedores do sistema, os resultados sobre o gênero podem sugerir que ferramentas de suporte com mais dados sejam fornecidas de forma resumida para o gênero masculino. Por outro lado, o gênero feminino é mais adequado relatórios detalhados.

Em relação ao grau de conhecimento, ficou evidente que esse aspecto direciona a atenção para as áreas de interesse que são mais importantes para o processo de elicitação

(Choice e Graph). Por esse motivo, pode-se sugerir ao analista de decisão que ao lidar com decisores com conhecimento menos suporte será requerido, ao contrário do que deve acontecer aos decisores sem conhecimento. Para o perfil de decisor sem conhecimento, o analista deverá voltar sua atenção para as instruções, fornecendo informações no início do processo para que esse grupo não divida sua atenção com elementos que causem distrações. Aos desenvolvedores do sistema, sugere-se uma alteração no layout da avaliação intracritério. Adicionando rodadas testes para familiarizar os usuários com o software, diminuindo a busca deles por elementos informativos. Contudo, vale salientar, que caso o decisor tenha já possui conhecimento ela tenha a opção de começar o experimento sem passar pela etapa de familiarização.

5.3. Limitações e Sugestões de Trabalhos Futuros

O presente estudo trouxe contribuições importantes, mas alguns aspectos merecem atenção para guiar trabalhos futuros. Um dos desafios enfrentados foi o tamanho e a composição da amostra: os 42 participantes foram divididos em quatro grupos (homens e mulheres, com e sem conhecimento), mas com tamanhos diferentes, o que pode ter introduzido alguma variabilidade nos resultados. Apesar disso, os dados coletados ofereceram uma base sólida para análises relevantes dentro do tema proposto.

Outro ponto foi o tempo significativo dedicado à limpeza dos dados gerados pelo eye-tracking, um processo essencial para garantir a qualidade das análises, mas que demandou alto esforço. Esse fator pode limitar a escalabilidade do método, especialmente em estudos maiores ou com volumes maiores de dados.

Embora o experimento tenha utilizado tanto o eye-tracking quanto o EEG como ferramentas de análise comportamental, o foco deste trabalho foi apenas na análise dos dados de eye-tracking. Isso deixou uma grande oportunidade para investigar como os padrões visuais se relacionam com a atividade neural capturada pelo EEG, o que poderia trazer uma compreensão mais profunda dos processos cognitivos e perceptivos analisados.

Para trabalhos futuros, seria interessante equilibrar os tamanhos dos grupos e ampliar a amostra, buscando reduzir possíveis variações e aumentar a generalização dos resultados. Além disso, investir em técnicas que automatizem o pré-processamento dos dados pode tornar a análise mais eficiente e viável. Outro avanço seria integrar os dados de EEG nas análises, explorando conexões entre o comportamento visual e a atividade cerebral. Por fim, replicar o experimento em diferentes contextos e populações poderia ampliar o alcance e a aplicação

prática dos resultados, contribuindo significativamente para a evolução das pesquisas nessa área.

Outra oportunidade relevante para trabalhos futuros, é a integração de técnicas de deep learning e aprendizagem de máquina para classificar o decisor quanto as características de gênero e grau de conhecimento. A criação desse modelo poderia oferecer ao SAD do FITradeoff uma adaptação automática e em tempo real, conforme o modelo classificasse o decisor a interface seria modificada de forma que as perdas de atenção e inconsistência fossem reduzidas, tornando o software mais eficiente e eficaz.

REFERÊNCIAS

- Pan, B., & Zhong, Y. (2011). *Eye Tracking and Web Usability: Understanding How Users Interact with Websites*. *International Journal of Human-Computer Interaction*, 27(5), 467-479.
- Rayner, K. (1998). *Eye movements in reading and information processing: 20 years of research*. *Psychological Bulletin*, 124(3), 372-422.
- Newell, A., & Rosenbloom, P. S. (1981). The Power Law of Practice. In J. R. Anderson (Ed.), *Cognitive Skills and Their Acquisition* (pp. 1–55). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics* (6th ed.). Pearson Education.
- Sokal, R. R., & Rohlf, F. J. (1995). *Biometry: The Principles and Practice of Statistics in Biological Research* (3rd ed.). W. H. Freeman and Co.
- Howell, D. C. (2012). *Statistical Methods for Psychology* (8th ed.). Cengage Learning.
- Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2005). *Applied Linear Statistical Models* (5th ed.). McGraw-Hill.
- Hinkle, D. E., Wiersma, W., & Jurs, S. G. (2003). *Applied Statistics for the Behavioral Sciences* (5th ed.). Houghton Mifflin.
- Šidák, ZK (1967). "Regiões de confiança retangulares para as médias de distribuições normais multivariadas". *Journal of the American Statistical Association* . 62 (318): 626–633. doi : 10.1080/01621459.1967.10482935 .
- Mauchly, J. W. (1940). «Significance Test for Sphericity of a Normal n-Variate Distribution». *The Annals of Mathematical Statistics*. 11 (2): 204–209. JSTOR 2235878. doi:10.1214/aoms/1177731915
- Greenhouse, S. W.; Geisser, S. (1959). "On methods in the analysis of profile data". *Psychometrika*. 24: 95–112.
- Montgomery, D. C. (2017). *Design and Analysis of Experiments*. Wiley.
- Fisher, R. A. (1925). *Statistical Methods for Research Workers*.
- Tukey, J. W. (1953). *The problem of multiple comparisons*. *American Statistical Association Journal*, 58(3), 35-41.
- Cutler, D. R., Edwards Jr, T. C., Beard, K. H., Hess, K. T., Gibson, J., & Lawler, J. J. (2007). Random forests for classification in ecology. *Ecology*, 88(11), 2783-2792. <https://doi.org/10.1890/07-0539.1>

- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32.
<https://doi.org/10.1023/A:1010933404324>
- Chen, T., & Guestrin, C. (2016). *XGBoost: A scalable tree boosting system*. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 785-794).
- Rice, J. A. (2007). *Mathematical Statistics and Data Analysis* (3rd ed.). Belmont, CA: Thomson Brooks/Cole.
- Casella, G., & Berger, R. L. (2002). *Statistical Inference* (2nd ed.). Duxbury.
- DeGroot, M. H., & Schervish, M. J. (2012). *Probability and Statistics* (4th ed.). Pearson Education.
- Burez, J., & Van den Poel, D. (2009). *Handling class imbalance in customer churn prediction*. *Expert Systems with Applications*, 36(3), 4626-4636.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer.
- Liaw, A., & Wiener, M. (2002). *Classification and regression by randomForest*. *R news*, 2(3), 18-22.
- Turban, E., Pollard, C., & Wood, G. (2018). *Information Technology for Management: On-Demand Strategies for Performance, Growth and Sustainability*. Wiley.
- Cohen, J., & Ranganath, K. (2007). *Neuroeconomics: How the Brain Shapes Decision-Making*. *Current Opinion in Neurobiology*, 17(2), 123-131.
- Zhang, P., & Li, N. (2005). The Influence of Gender Differences on Information Processing and Decision-Making in SAD Contexts. *Journal of Information Systems*, 10(3), 115–130.
- Hollender, M., & Gergely, G. (2018). Cognitive Adaptation and Gender Sensitivity in Decision-Making Systems. *Cognitive Systems Research*, 52, 57–65.

Damasio, A. R. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain*. New York: Penguin Books.

Wedel, M., & Pieters, R. (2008). *Eye Tracking for Visual Marketing. Foundations and Trends in Marketing*.

Luck, S. J. (2014). *An Introduction to the Event-Related Potential Technique*. MIT Press.

Ariely, D., & Berns, G. S. (2010). Neuromarketing: the hope and hype of neuroimaging in business. *Nature Reviews Neuroscience*.

Riedl, R., Mohr, P. N., Kenning, P., Davis, F. D., & Heekeren, H. R. (2010). Trusting Humans and Avatars: A Brain Imaging Study Based on Evolution Theory. *Journal of Management Information Systems*.

Fehr, E., & Camerer, C. F. (2007). Social neuroeconomics: the neural circuitry of social preferences. *Trends in Cognitive Sciences*.

Raichle, M. E., & Mintun, M. A. (2006). *Brain work and brain imaging. Annual Review of Neuroscience, 29(1)*, 283-304.

Plassmann, H., Ramsøy, T. Z., & Milosavljevic, M. (2012). Branding the brain: A critical review and outlook. *Journal of Consumer Psychology, 22(1)*, 18-36.

Lindstrom, M. (2008). *Buyology: Truth and Lies About Why We Buy*. Random House.

Da silva, Anderson Lucas Carneiro de Lima; Costa, Ana Paula Cabral Seixas; De Almeida, Adiel Teixeira. Exploring cognitive aspects of fittradeoff method using neuroscience tools. *Annals of operations research*, v. 312, p. 1147-1169, 2022.

ALAMRI, A. A.; BALKHI, Z. T. The effects of learning and forgetting on the optimal production lot size for deteriorating items with time varying demand and deterioration rates. *International Journal of Production Economics*, v. 107, n. 1, p. 125-138, 2007. <http://dx.doi.org/10.1016/j.ijpe.2006.08.004>

BADIRU, A. B. Computational Survey of Univariate and Multivariate Learning Curve Models. *IEEE Transactions on Engineering Management*, v. 39, n. 2, p. 176-188, 1992. <http://dx.doi.org/10.1109/17.141275>

BAI, J.; WANG, M.-Z.; WANG, J.-B. Single machine scheduling with a general exponential learning effect. *Applied Mathematical Modelling*, v. 36, n. 2, p. 829-835, 2012.

BROEK, M. V. *et al.* Effects of technological learning on future cost and performance of power plants with CO₂ capture. *Progress in Energy and Combustion Science*, v. 35, n. 6, p. 457-480, 2009.

CHATZIMICHALI, A. P.; TOURASSIS, V. D. An Empirical Investigation of Learning Curve Laws in the Tile Manufacturing Industry. In: *IEEE INTERNATIONAL CONFERENCE ON INDUSTRIAL ENGINEERING AND ENGINEERING MANAGEMENT*, 2007, Singapura. Proceeding... Singapura, 2007. p. 1110-1114.

CHIU, H. N.; CHEN, H. M. An optimal algorithm for solving the dynamic lot-sizing model with learning and forgetting in setups and production. *International Journal of Production Economics*, v. 95, n. 2, p. 179-193, 2005.

JABER, M. Y.; BONNEY, M. Production breaks and the learning curve: The forgetting phenomenon. *Applied Mathematical Modelling*, v. 20, n. 20, p. 162-169, 1996a.

ARES, G., MAWAD, F., GIMÉNEZ, A., & MAICHE, A. Influence of rational and intuitive thinking styles on food choice: Preliminary evidence from an eye-tracking study with yogurt labels. *Food Quality and Preference*, 31, 28-37, 2014.

BAULT, N.; WYDOODT, P., CORICELLI, G. Different Attentional Patterns for Regret and Disappointment: An Eye-tracking Study. *Journal of Behavioral Decision Making*, 29, 194-205, 2016.

BECHARA, A, DAMASIO, A.R. The somatic marker hypothesis: A neural theory of economic decision. *Games and economic behavior*, 52(2), 336-372, 2005.

BROCAS, I. Information processing and decision-making: evidence from the brain sciences and implications for Economics. *Journal of Economic Behavior & Organization*, 83(3), 292-310, 2012.

BUETTNER, R. Investigation of the relationship between visual website complexity and users' mental workload: A NeuroIS perspective. In: DAVIS, F.D., RIEDL, R., VOM BROCKE, J., LÉGER, P.-M., RANDOLPH, A. (Eds.). *Information Systems and Neuroscience*, Berlin: Springer, 2015

CHOI, J. J., LAIBSON, D., MADRIAN, B. C., & METRICK, A. Reinforcement learning and savings behavior. *The Journal of finance*, 64(6), 2515-2534, 2009.

CHUANG, H; LIN, C; CHEN, Y. Exploring the triple reciprocity nature of organizational value cocreation behavior using multicriteria decision making analysis. *Mathematical Problems in Engineering*, 1-15, 2015.

Castells, M. (1996). *The Rise of the Network Society*. Blackwell Publishers.

Brynjolfsson, E., & McAfee, A. (2014). *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*. W. W. Norton & Company

Mayer-Schönberger, V., & Cukier, K. (2013). *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. Houghton Mifflin Harcourt.

de Almeida, A.T., Frej, E.A. & Roselli, L.R.P. Combining holistic and decomposition paradigms in preference modeling with the flexibility of FITradeoff. *Cent Eur J Oper Res* 29, 7–47 (2021).

de Almeida, J. A. A., Frej, E.A. & Roselli, L.R.P., de Almeida, A.T. 2024. Analytical aspects of combining holistic evaluation and decomposition elicitation for preference modeling in the FITradeoff method. *International Transactions in Operational Research*. 1-42. <http://doi.org/10.1111/itor.13470>.

de Almeida, J. A. A., Frej, E.A. & Roselli, L.R.P., de Almeida, A.T. 2024. Analytical aspects of combining holistic evaluation and decomposition elicitation for preference modeling in the FITradeoff method. *International Transactions in Operational Research*. 1-42. <http://doi.org/10.1111/itor.13470>.

DE ALMEIDA AT, FREJ EA, ROSELLI LRP & COSTA APCS. 2023. A summary on FITradeoff method with methodological and practical developments and future perspectives. *Pesquisa Operacional*, 43:e268356. doi: 10.1590/0101-7438.2023.043spe1.00268356.