

Controlled Experiment with Users Report about Discord

André Monteiro
fc51718

Faculty of Sciences, University of
Lisbon
Lisbon, Portugal
fc51718@alunos.fc.ul.pt

Frederico Vilante
fc49019

Faculty of Sciences, University of
Lisbon
Lisbon, Portugal
fc49019@alunos.fc.ul.pt

Rodrigo Albino
fc49027

Faculty of Sciences, University of
Lisbon
Lisbon, Portugal
fc49027@alunos.fc.ul.pt

SYSTEM

O sistema envolvido na experiência foi o Discord, uma plataforma de comunicação social baseada, principalmente, em servidores comunitários. Esta plataforma tem vindo a crescer de forma exponencial como o meio primordial de comunicação pelos membros das faixas etárias mais novas. Além de facilitar a comunicação, o Discord permite que indivíduos se juntem e criem os servidores comunitários, mencionados acima, para falarem de temas de interesse comum.

RESEARCH QUESTION

RQ1: Será que há efeito do género do utilizador sobre os tempos totais de conclusão das tarefas dos servidores comunitários do Discord?

RQ2: Será que é mais fácil procurar por um servidor comunitário público fora do Discord, em sites externos, do que dentro do Discord utilizando a funcionalidade de explorar servidores públicos?

USER STUDY DESIGN

Este estudo com utilizadores teve como objetivo recolher dados para entendermos se as funcionalidades do Discord mais concretamente, em servidores comunitários, diferem consoante o género e se a funcionalidade de procura de servidores comunitários dentro do discord suporta as necessidades dos utilizadores. Além disso, quisemos entender se realmente existe uma escassez de funcionalidades na procura de servidores dentro do discord ou não.

Para isso decidimos fazer um ensaio controlado onde os participantes têm que fazer diversas tarefas que propomos e responder da maneira mais transparente possível. Além dessas tarefas também adicionamos ao guião um Single Easy Question (SEQ) que serve para entender a facilidade que teve nas tarefas, um User Experience Questionnaire (UEQ) para podermos entender a sua experiência de utilização e finalmente fizemos perguntas para percebermos a frequência de utilização que cada participante tem nas tarefas que irá realizar e perguntas para fazermos a sua caracterização demográfica. Durante as recolhas dos dados que efetuamos foram recolhidas gravações de vídeo e captura do ecrã com a devida autorização dos participantes.

Em simultâneo com a elaboração do Guião do Procedimento com as tarefas a serem realizadas pelos nossos utilizadores, fomos elaborando os Task Scenarios para cada tarefa, os quais apresentamos em seguida, por ordem de realização das tarefas. Cada Task Scenario é composto pelo contexto e pela tarefa a ser realizada.

O grupo decidiu avançar com um conjunto de tarefas que acabaram por trazer maior parte das funcionalidades mais importantes dos servidores comunitários.

Para tal, usamos as seguintes tarefas, ordenadas por ordem de realização:

PROTOCOLO DO PROCEDIMENTO EXPERIMENTAL FINAL: TAREFAS E TASK SCENARIOS

Task 1

Task Scenario: Na última semana, tu e os teus amigos, que têm contas no Discord, têm vindo a conversar sobre temas atuais que interessam a todos mas têm restrições de tempo para poderem estarem juntos presencialmente. Os teus amigos pediram-te para arranjares uma forma de no Discord poderem, por um lado, continuar as suas conversas e, por outro lado, poderem ouvir música, manter conversas separadas por assuntos e jogar online enquanto convivem. Utiliza as funcionalidades do Discord para estabeleceres uma forma de poderes jogar online enquanto manténs conversas separadas por assuntos com os teus amigos.

Task: Adiciona um novo Servidor, tanto para ti como para os teus amigos, com o nome "servidor2000".

Task 2

Task Scenario: No último mês, o servidor que criaste com os teus amigos tem vindo a crescer, e tu que és administrador e vinte amigos teus que são moderadores, repararam que a gestão está a ficar demasiado confusa para o vosso gosto, porque os membros quando precisam de conversar sobre temas gerais fazem-no dentro dos canais específicos devido à falta de um canal abrangente. Os moderadores do servidor pediram-te que arranjaasses uma forma de poderem escrever e falar com os membros para poderem responder e ajudar com as suas preocupações gerais. Usa as funcionalidades disponíveis dentro do servidor que criaste para o organizar de modo que todos os membros tenham uma parte onde possam escrever e falar sobre assuntos gerais sem se misturarem com os assuntos específicos do servidor.

Task: Entra no servidor que criaste e cria dois canais novos, um de texto e outro de voz, ambos com o nome "geral".

Task 3

Task Scenario: Nos últimos dois dias, tu e os teus amigos decidiram ir mergulhar num recife de corais. Num canal do teu servidor, tens estado a conversar sobre câmaras para tirarem fotos a peixes ao redor do recife de coral, sendo que têm um orçamento limitado. Tu não sabes quantos megapixéis são importantes para terem uma boa qualidade de imagem, mas tens em mente que um dos teus amigos do Discord o “MEU2022#4061”, que não está no servidor que criaste, é especialista em câmaras subaquáticas e é mergulhador profissional. Utiliza as funcionalidades do Discord para introduzires este teu amigo “MEU2022#4061” na conversa do canal para que tu e os teus amigos possam encontrar uma câmara subaquática adequada dentro do orçamento disponível.

Task: Convida para o servidor que acabaste de criar o usuário com o nome “MEU2022#4061”.

Task 4

Task Scenario: Nos últimos meses, notaste que muitos dos novos membros que aderem ao canal ficam intimidados a interagir e participar nos primeiros tempos após a adesão ao servidor. Alguns membros explicaram que nos primeiros tempos se sentem pouco integrados por estarem pouco familiarizados. Escreve uma mensagem geral de notificação acolhedora que esteja visível para todos os atuais e futuros membros que venham a aderir ao teu servidor para que os novos membros sejam incentivados a interagir e participar e para que se sintam integrados.

Task: No canal de texto “geral” que tu criaste envia a tua primeira mensagem com o conteúdo “@everyone Sê bem-vindo ao servidor 2000. Fica à vontade para explorar e conviver.” e fixa esta mensagem.

Task 5

Task Scenario: Na última semana, notaste que as notificações do servidor que criaste no Discord, nos teus dispositivos, como smartphone, computador e tablet, estavam a ser muito frequentes para o teu gosto e que a maioria eram sobre assuntos que te eram indiferentes. Utiliza as opções do servidor que criaste para limitar as notificações do sistema, para que apenas sejas notificado quando o teu nome de utilizador for mencionado nas mensagens deste servidor.

Task: Muda o conjunto de notificações do servidor que criaste para a opção “Only @mentions”.

Task 6

Task Scenario: Estás a planear participar num concurso de tecnologia com os teus amigos do servidor 2000 com um robô para ajudar a dar resposta a uma emergência atual que ocorreu no vosso país e que afeta todos os membros do servidor. O projeto do robô é do interesse de todos os membros do servidor. Haverá um projeto para o robô que vão construir que tem de ser entregue à entidade organizadora do concurso, no qual tens estado a colaborar. Os teus amigos pediram para partilhasses o projeto e que pudessem ir falando e dando opiniões para poderem melhorar o projeto do robô. Usa as funcionalidades do Discord para dentro do servidor 2000 poderes mostrar ao vivo o projeto do robô em que tens estado a colaborar e para, em simultâneo, com os teus amigos poderem falar para trocarem ideias sobre o robô.

Task: No servidor que criaste entra no canal de voz “geral” e seleciona partilhar o teu ecrã principal.

Task 7

Task Scenario: No último ano, começaste a interessar-te por um determinado tema e tornaste-te um profissional desse tema. Entretanto soubeste que várias pessoas, que também estão no Discord, trocam ideias e falam entre si sobre esse tema. Apesar de não saberes onde te dirigires para encontrar esses locais de conversa online, tens ideia como o Discord está organizado e estás muito interessado em conviver com outras pessoas interessadas nesse tema. Utiliza os recursos disponíveis no Discord para encontrares um servidor público onde possas conversar e trocar ideias com outros utilizadores e onde possas aprender mais sobre esse tema específico.

Task: Procura um servidor, na funcionalidade “explore public servers” no Discord, com os seguintes requisitos:

- Tema: Música;
- Número de users: entre 100 a 1000;
- Com a descrição: “This is a server for all fellow alternative music lovers, whether you like indie, new-wave or whatever this is the community for you!”.

Task 8

Task Scenario: Nos últimos dois meses, tu conheceste uma pessoa que é profissional de um determinado tema com a qual tens vindo a trocar ideias e desde então ficaste interessado em aprofundar o teu conhecimento sobre esse tema. Tu sabes como o Discord está organizado e estás muito interessado em conversar com outras pessoas sobre esse tema. Entretanto outro amigo teu partilhou contigo um site exterior ao Discord com listas de servidores comunitários onde se pode pesquisar sobre um tema específico. Utiliza as funcionalidades disponíveis neste site exterior para encontrares um servidor público do Discord sobre esse tema que te interessa para poderes conversar com outros utilizadores, trocares ideias e aprenderes mais sobre esse tema.

Task: Procura um servidor, agora num site externo de pesquisa de servidores (<https://top.gg/servers>), que já está aberto no computador, com os seguintes requisitos:

- Tema: Música;
- Número de users: entre 100 a 1000;
- Com a descrição: “This is a server for all fellow alternative music lovers, whether you like indie, new-wave or whatever this is the community for you!”.

DESCRIÇÃO DO PROTOCOLO DO PROCEDIMENTO EXPERIMENTAL UTILIZADO

Além das tarefas, foi decidido pedirmos a utilizadores, com vários graus de experiência de utilização com o sistema, para demonstrarem a sua aptidão para se obter um maior leque de informação possível e que englobasse todos os tipos de utilizadores existentes no sistema. Com esta ideia, foi decidido usar o within-subjects study.

Para realizar a experiência controlada, escolhemos locais físicos para a realização presencial das experiências controladas com os utilizadores. Esta experiência controlada foi feita de forma anónima respeitando a privacidade das pessoas que respondem ao mesmo, as imagens que foram filmadas não foram partilhadas e depois da utilização das mesmas para produzir os dados necessários não foram usadas para outros fins. Além disso, solicitamos o seu consentimento para usar os seus dados anónimos para o nosso estudo e informamos que eles podem desistir a qualquer momento da experiência controlada e retiraremos seus dados anónimos. Depois disso, eles só precisaram fazer as tarefas e responder às perguntas e entregar o guião preenchido. Além disso, ao longo da execução das tarefas optamos por gravar em vídeo as mãos dos nossos participantes a completar as tarefas para que depois pudéssemos confirmar a contagem do número de cliques no mouse, para além de termos também gravado a captura da tela do computador enquanto os participantes faziam as tarefas para percebermos os passos dados para conseguirem concluir as tarefas. Foram consideradas duas variáveis: uma dependente, o tempo; e uma independente, o género do utilizador.

Assim, nós pretendemos recolher informação sobre o tempo que cada utilizador leva a fazer cada tarefa que é pedida, qual foi a facilidade dos utilizadores de fazerem as tarefas, se os resultados estão diretamente ligados com o género do utilizador, e a diferença entre a procura externa e no próprio Discord.

Para conseguirmos ter uma boa comparação entre os géneros, organizamos os dados em dois grupos, o grupo Masculino e o grupo Feminino, e fizemos testes a utilizadores com idades entre os 18 e os 29 anos, pois nos questionários que fizemos anteriormente conseguimos entender que a idade acabava por influenciar nos tempos. Todo o estudo foi feito com máquinas com o poder suficiente para correr a plataforma sem qualquer erro e sem qualquer problema de lag nem nada do tipo. Junto com o dispositivo foi usado um mouse que continha exatamente os mesmos “dots per linear inch” (DPI) e “polling rate” que é medido em Hz. Além disso, todos os estudos foram feitos num ambiente confortável e propício para que não houvesse qualquer distração e fosse feita a experiência controlada com as melhores condições possíveis. Esta experiência de estudos com utilizadores, apesar de não ter um tempo obrigatório, em nenhum dos casos ultrapassou os 20 minutos aproximadamente na sua execução.

Deste modo, iniciamos a recolha dos dados com um estudo piloto com duas pessoas, uma pertencente ao género masculino e a outra pertencente ao género feminino, para entendermos se haviam alguns aspetos errados no guião do procedimento e o que precisava de ser melhorado tanto nas tarefas que propusemos como na escrita de todo o guião.

Depois destes ajustes com o teste piloto avançámos para a experiência controlada com 12 pessoas propriamente dita, sendo que tal como era o nosso objetivo a priori, obtivemos no final uma proporção igualitária do género feminino e masculino que nos permitiu ter o melhor conjunto de dados possíveis representativos dos dois grupos Masculino e Feminino.

SINGLE EASE QUESTION

O primeiro teste HCI Survey que foi pedido que os participantes completassem foi o Single Ease Question (SEQ), o qual dá uma medida da facilidade. Neste teste nós aplicamos a seguir a cada tarefa, e perguntamos se tinha sido fácil essa tarefa em específico, de forma a percebermos a facilidade de utilização de cada utilizador em relação a cada tarefa. A pergunta do Single Ease Question (SEQ) que aplicamos no guião do procedimento é apresentada em seguida.

Em termos gerais, esta tarefa foi? <small>(assinala com um X uma opção)</small>							Single Ease Question (SEQ)
Muito difícil						Muito fácil	
1	2	3	4	5	6	7	

Tabela 1. Tabela exemplo do Single Ease Question

USER EXPERIENCE QUESTIONNAIRE

O segundo teste HCI Survey que pedimos aos nossos participantes no estudo que respondessem foi o User Experience Questionnaire (UEQ). Com este HCI Survey pretendemos perceber a experiência de utilização, sendo que o mesmo foi aplicado apenas no final de todas as tarefas terem sido concluídas. As perguntas do User Experience Questionnaire (UEQ) que aplicamos no guião do procedimento são apresentadas em seguida.

User Experience Questionnaire (UEQ) Experiência de utilização das tarefas (assinale com um X uma opção na escala de 1 a 7)								
Aborrecido								Excitante
Mau								Bom
Desagradável								Agradável
Insuportável								Prazeroso
Pouco Atraente								Muito Atraente
Nao Amigável								Amigável
Lento								Rápido
Ineficiente								Eficiente
Imprático								Prático
Desorganizado								Organizado
Imprevisível								Previsível
Obstrui								Apoia
Inseguro								Seguro
Não satisfaz expectativas								Satisfaz expectativas

Tabela 2. Tabela UEQ usada

CARACTERIZAÇÃO DEMOGRÁFICA DOS PARTICIPANTES NA EXPERIÊNCIA CONTROLADA

Finalmente após terem sido feitas as tarefas do experimento, preenchido o Single Ease Question (SEQ) e o User Experience Questionnaire (UEQ), fizemos uma parte final de perguntas para percebermos a frequência habitual de utilização que tem nas tarefas que irá realizar, numa escala de Likert, e perguntas para fazermos a sua caracterização demográfica. Esta parte final do guião do procedimento inclui para além das perguntas da frequência de utilização habitual das tarefas pelos participantes, perguntas sobre a idade, o género a que pertence, o país onde residia, a ocupação profissional, e o nível de escolaridade.

TESTES PILOTO

A fim de validarmos o nosso método experimental conduzimos o teste piloto. Com base nos resultados obtidos com o teste piloto com as duas pessoas, alterámos a escrita de algumas perguntas para se perceber melhor. Também concluímos que uma das nossas tarefas iniciais tinha uma taxa de sucesso praticamente nula. Nessa tarefa era pedido ao participante que criasse uma role chamada “teste123” e que adicionasse essa role ao usuário que tinham convidado anteriormente. Como considerámos que essa pergunta estava muito fora das componentes principais do estudo relativas aos servidores comunitários do Discord, optámos por excluir essa tarefa do guião do procedimento. Para além disso, na tabela em que perguntávamos aos nossos participantes a sua frequência habitual de realização daquelas tarefas no seu quotidiano, optámos por alargar o tempo analisado do último mês para os últimos 6 meses, pois precisávamos de um range maior de tempo.

RESULTS

STATISTICAL ANALYSIS

Para descrevermos a amostra fizemos uma análise descritiva onde apresentamos as médias, desvios padrões, variâncias, entre outras do SEQ, tempo e dos cliques por tarefa. Além disso, fizemos também para as idades.

	Tarefa 1				Tarefa 2			Tarefa 3			Tarefa 4			Tarefa 5			Tarefa 6		
Measures of central tendency	Idade	SEQ	tempo (s)	click	SEQ	tempo (s)	click	SEQ	tempo (s)	click	SEQ	tempo (s)	click	SEQ	tempo (s)	click	SEQ	tempo (s)	click
Mean	22,00	6,17	32,67	8,50	6,25	32,42	10,25	5,17	41,75	8,83	6,25	37,58	6,17	4,92	16,25	7,08	6,17	24,25	7,17
Median	22,00	6,50	25,00	5,50	6,00	23,00	7,50	6,00	35,00	7,00	6,50	40,00	4,50	6,00	9,50	5,50	6,00	21,00	6,50
Mode	23,00	7,00	23,00	5,00	7,00	#N/D	4,00	6,00	#N/D	5,00	7,00	#N/D	4,00	6,00	8,00	2,00	7,00	21,00	9,00
Measures of spread																			
Range	5,00	4,00	77,00	25,00	2,00	79,00	25,00	5,00	115,00	17,00	3,00	62,00	12,00	6,00	50,00	18,00	3,00	47,00	12,00
Variance	2,00	1,42	452,61	49,55	0,57	654,08	52,20	3,06	1098,39	30,33	0,93	408,08	13,61	3,90	206,02	31,17	0,88	185,11	11,06
Standard deviations	1,41	1,19	21,27	7,04	0,75	25,58	7,23	1,75	33,14	5,51	0,97	20,20	3,69	1,98	14,35	5,58	0,94	13,61	3,33
Q1 (First Quartile)	21,75	6,00	20,75	5,00	6,00	17,50	5,00	3,75	17,75	5,00	6,00	18,50	4,00	4,75	6,75	2,75	6,00	15,50	4,75
Q3 (Third Quartile)	23,00	7,00	40,00	8,75	7,00	31,75	13,50	6,25	53,50	11,75	7,00	52,75	6,75	6,00	21,75	10,25	7,00	32,25	9,00
Interquartile range (IQR)	1,25	1,00	19,25	3,75	1,00	14,25	8,50	2,50	35,75	6,75	1,00	34,25	2,75	1,25	15,00	7,50	1,00	16,75	4,25

Tabela 3. Tabela geral de todas as tarefas com médias, desvios padrões, entre outras

Para entendermos melhor quais são as diferenças entre os dois géneros decidimos dividir e fazer a análise também separada.

		Tarefa 1			Tarefa 2			Tarefa 3			Tarefa 4			Tarefa 5			Tarefa 6		
Measures of central tendency	Idade	SEQ	tempo (s)	click	SEQ	tempo (s)	click	SEQ	tempo (s)	click	SEQ	tempo (s)	click	SEQ	tempo (s)	click	SEQ	tempo (s)	click
Mean	22.17	6.83	21.00	5.67	6.50	18.67	6.33	6.00	27.00	6.50	6.50	33.50	5.17	5.67	12.83	5.17	6.50	18.17	6.50
Median	22.50	7.00	20.00	5.00	6.50	18.50	6.00	6.00	17.50	5.50	6.50	34.50	4.50	6.00	8.00	4.00	6.50	17.00	7.00
Mode	23.00	7.00	#N/D	5.00	7.00	#N/D	4.00	6.00	#N/D	5.00	7.00	#N/D	4.00	6.00	8.00	3.00	7.00	#N/D	9.00
Measures of spread																			
Range	5.00	1.00	18.00	4.00	1.00	14.00	6.00	3.00	52.00	7.00	1.00	41.00	6.00	3.00	28.00	8.00	1.00	26.00	6.00
Variance	2.97	0.17	51.20	2.27	0.30	23.47	5.87	1.20	442.00	6.70	0.30	195.50	4.57	1.07	117.37	10.17	0.30	74.97	6.70
Standard deviations	1.72	0.41	7.16	1.51	0.55	4.84	2.42	1.10	21.02	2.59	0.55	13.98	2.14	1.03	10.83	3.19	0.55	8.66	2.59
Q1 (First Quartile)	22.00	7.00	15.50	5.00	6.00	15.75	4.25	6.00	13.25	5.00	6.00	28.50	4.00	5.25	6.50	3.00	6.00	14.50	4.50
Q3 (Third Quartile)	23.00	7.00	26.00	6.50	7.00	19.75	7.75	6.75	38.25	7.50	7.00	40.50	5.75	6.00	17.75	7.25	7.00	20.25	8.75
Interquartile range (IQR)	1.00	0.00	10.50	1.50	1.00	4.00	3.50	0.75	25.00	2.50	1.00	12.00	1.75	0.75	11.25	4.25	1.00	5.75	4.25

Tabela 4. Tabela de todas as tarefas com médias, desvios padrões, entre outras para o género masculino

		Tarefa 1			Tarefa 2			Tarefa 3			Tarefa 4			Tarefa 5			Tarefa 6		
Measures of central tendency		SEQ	tempo (s)	click	SEQ	tempo (s)	click	SEQ	tempo (s)	click	SEQ	tempo (s)	click	SEQ	tempo (s)	click	SEQ	tempo (s)	click
Mean	21.83	5.50	44.33	11.33	6.00	46.17	14.17	4.00	56.50	11.17	6.00	43.50	7.17	4.17	44.67	9.00	5.83	30.33	7.83
Median	22.00	6.00	42.00	8.50	6.00	32.50	14.00	3.50	43.50	11.00	6.50	50.50	5.00	5.50	22.00	8.50	6.00	29.00	6.50
Mode	23.00	6.00	#N/D	#N/D	7.00	#N/D	#N/D	3.00	#N/D	#N/D	7.00	#N/D	4.00	6.00	#N/D	2.00	6.00	#N/D	#N/D
Measures of spread																			
Range	3.00	4.00	68.00	25.00	2.00	76.00	24.00	5.00	105.00	17.00	3.00	61.00	12.00	5.00	178.00	18.00	3.00	45.00	11.00
Variance	1.37	1.90	617.87	87.47	0.80	961.77	72.17	3.20	1452.30	46.97	1.60	582.70	22.97	6.17	4681.87	49.60	1.37	243.47	16.57
Standard deviations	1.17	1.38	24.86	9.35	0.89	31.01	8.50	1.79	38.11	6.85	1.26	24.14	4.79	2.48	68.42	7.04	1.17	15.60	4.07
Q1 (First Quartile)	21.25	5.25	26.75	5.25	5.25	27.25	8.50	3.00	34.50	5.75	5.25	24.25	4.00	2.00	10.25	3.00	5.25	22.25	5.25
Q3 (Third Quartile)	22.75	6.00	46.75	12.50	6.75	67.00	15.75	4.75	69.00	16.25	7.00	57.25	9.75	6.00	27.75	12.50	6.75	38.00	9.25
Interquartile range (IQR)	1.50	0.75	20.00	7.25	1.50	39.75	7.25	1.75	34.50	10.50	1.75	33.00	5.75	4.00	17.50	9.50	1.50	15.75	4.00

Tabela 5. Tabela de todas as tarefas com médias, desvios padrões, entre outras para o género feminino

Conseguimos entender facilmente que a média de tempos das várias tarefas se encontra maior no género feminino e o SEQ se encontra mais baixo, o que diz que tiveram mais dificuldade em realizar as tarefas que o género masculino.

Temos uma média de idades de 22 anos, o que ficou entre o espaço que queríamos, sendo que o género masculino tem ligeiramente uma maior media que o género feminino. Vemos também que as medias de facilidade são um pouco diferentes entre os géneros, iremos fazer teste estatísticos para entender se essa diferença acaba por ser o suficiente.

Para fazermos realmente essas comparações entre géneros para conseguirmos realmente entender se estas diferenças são realmente significativas, fizemos testes estatísticos para o SEQ, como já referido anteriormente e o tempo.

Por os dados do SEQ serem likert scale, logo são dados ordinais, além disso os grupos são independentes entre si pois estão divididos por género e todas as pessoas fizeram exatamente as mesmas tarefas. Por isso o teste que fizemos, neste caso, foi um teste não paramétrico, Mann-Whitney U test. Decidimos fazer este teste por tarefa pois achamos que seria importante entender se realmente a diferença da facilidade de interseção entre os géneros se encontra em todas as tarefas. Em contrapartida em relação ao tempo decidimos fazer apenas o total no tempo e não por tarefa para sabermos um contexto global, sabendo que as nossas perguntas englobam grande parte das funcionalidades dos servidores comunitários e bastante importante fazer o tempo total pois acaba por nos dar dados de comparação entre géneros de um funcionamento geral dentro dos servidores. Todos os testes e outliers que foram feitos na aplicação R-studio.

De maneira a ver se existiam ou não outliers para o tempo no total como para os clicks totais, fizemos boxplots para ambos. Podemos ver que por exemplo no tempo existe 1 outlier que acaba por ser o valor de 351 segundos.

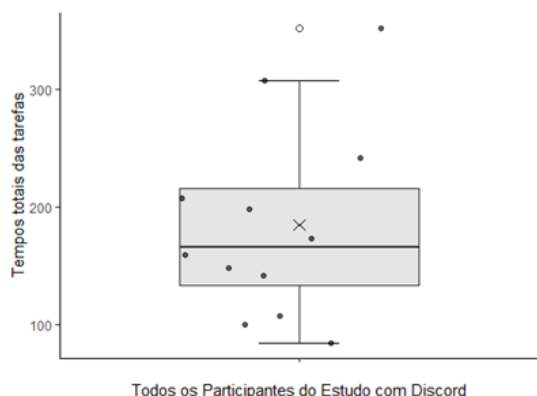


Figura 1. Box Plots do tempo total das tarefas

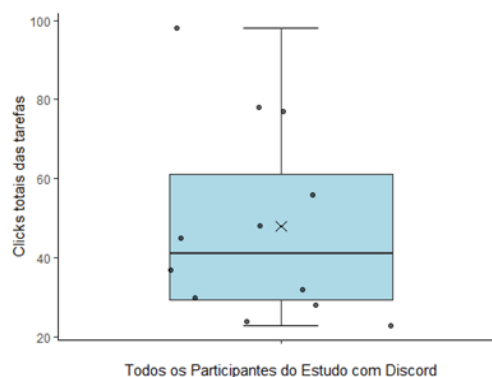


Figura 2. Box Plots dos clicks totais das tarefas

Fazendo então o Mann-Whitney U test sendo a hipótese nula que “the sum of the rankings in the two groups does not differ” e a hipótese alternativa “the sum of the rankings differs in the two groups”, que apenas a tarefa 1 rejeita a hipótese nula, todas

as outras aceitam pois os valores de p-value são maiores que 0,05 (Tarefa 1 p-value = 0.02922, Tarefa 2 p-value = 0.3413, Tarefa 3 p-value = 0.08584, Tarefa 4 p-value = 0.6628, Tarefa 5 p-value = 0.3452, Tarefa 6 p-value = 0.3423).

Antes de fazermos os testes estatísticos entre gêneros quisemos entender se a frequência de utilização se relacionava com o tempo que cada usuário que fez o experimento, para conseguirmos ter essas correlações fizemos o Pearson's correlation coefficient. Conseguimos entender que a tarefa 1 ($r = -0.254$), a tarefa 3 ($r = -0.457$), a tarefa 5 ($r = -0.514$) e a tarefa 6 ($r = -0.316$) têm uma baixa correlação negativa, enquanto a tarefa 2 ($r = -0.008$) e 4 ($r = 0.147$) não têm correlação.

Em relação ao melhor teste estatístico para ser feito para os dados do tempo entre gêneros, primeiro tivemos de entender se estes dados têm uma distribuição normal, além de que os grupos têm de apresentar variâncias que fossem homogêneas.

Para vermos que existe uma distribuição normal nos dados fizemos um teste Shapiro-Wilk Test, onde a hipótese nula será que os dados são normais e a hipótese alternativa que os dados não são normais. Neste teste o p-value deu o valor 0.3594, e por este valor ser maior que 0.05, vamos ficar com a hipótese nula, e vamos aceitar que o pressuposto da normalidade necessário para o Teste t-Student foi cumprido.

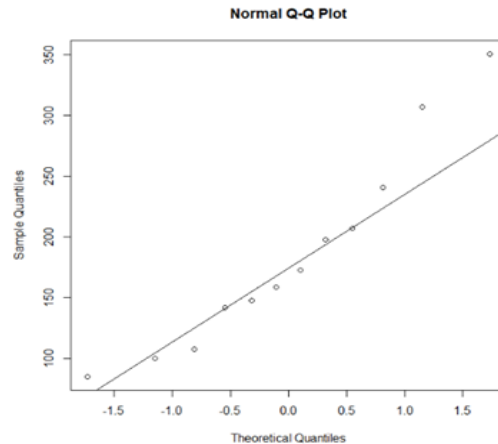


Figura 3. Gráfico dos valores do tempo total e linha de normalização

Caso os grupos não apresentassem variâncias que fossem homogêneas, nós ainda conseguiríamos usar o Teste t de Student independente, mas iríamos ter de fazer a Correção de Welch. Outro aspeto é que as nossas observações têm de ser independentes, sendo que esse pressuposto nós já sabíamos que tinha sido atendido porque isso dependia do nosso desenho experimental. Além disso, também não deveria haver outliers dentro dos grupos, porque se houvesse muitos outliers isso poderia afetar muito a média dos grupos e isso iria afetar o desenho do teste. Como já mostrado anteriormente podemos verificar que não foram encontrados muitos outliers, nem para o grupo "Masculino", nem para o grupo "Feminino".

E assim o que nos faltava era verificar se as variâncias eram homogêneas, pelo que decidimos fazer o teste de homogeneidade de variâncias, o Teste de Levene. E assim obtivemos o valor do p do Teste de Levene, que calculámos com base na média dos tempos totais, porque estávamos a fazer uma estatística paramétrica que tem como base a média. Para este teste de Levene tínhamos como hipótese nula que as variâncias dos dois grupos eram homogêneas e como hipótese alternativa que as variâncias dos dois grupos não eram homogêneas. Neste caso, como o nosso valor de p foi maior do que 0,05 (p-value = 0.174), então podemos considerar que as variâncias são homogêneas.

Pressupostos para que pudéssemos fazer o Teste t de Student:

1. Grupos com distribuição normal: Pressuposto atendido, verificado com Teste de Shapiro-Wilk
2. Homogeneidade de variâncias: Pressuposto atendido, verificado com Teste de Levene
Opção: correção de Welch
3. Observações independentes: Pressuposto atendido, já sabíamos a priori
4. Ausência de muitos outliers: Pressuposto atendido, verificado com o Box Plot

Depois dos passos todos mostrados anteriormente fomos finalmente analisar os valores do Teste t de Student para amostras independentes com hipóteses bicaudais, onde obtivemos os valores do t estatístico (t-stat), dos graus de liberdade (df) e o valor de p (p-value). Então o Teste t de Student tem como hipótese nula que os tempos entre os gêneros são iguais e como hipótese alternativa que os tempos dos dois grupos são diferentes.

O Teste t de Student para amostras independentes mostrou que há efeito do género sobre os tempos totais de conclusão das tarefas ($t_{(10)}=3,621$; $p=0,004681 \rightarrow p<0,05$).

Este p-value rejeita a hipótese nula e aceita a hipótese alternativa, que basicamente diz que há efeito do género sobre os tempos totais de realização das tarefas e que os grupos diferentes apresentam tempos diferentes.

MAIN FINDINGS AND DISCUSSION

Em relação à experiência de utilização se formos entender os resultados num contexto geral pelas escalas usadas entendemos que a experiência de utilização dos utilizadores acaba por não ser a melhor quando são feitos estes tipos de tarefas em servidores comunitários, ficando tudo abaixo da média. Apesar da amostra não ser a mais expressiva, isto pode indicar que poderá ser preciso uma adaptação destas tarefas para que os usuários comecem a ter uma melhor experiência nestas tarefas.

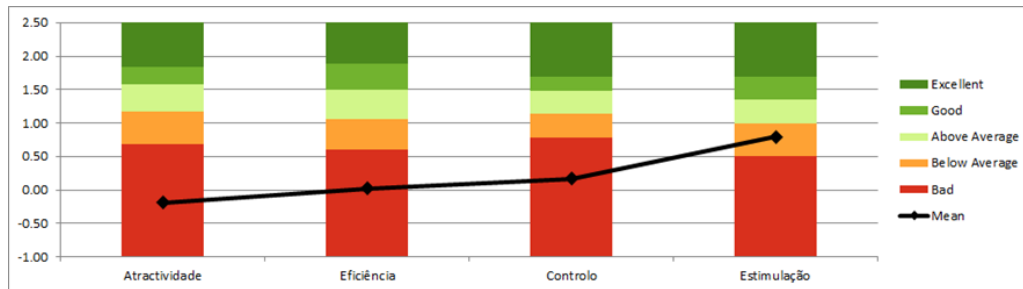


Figura 4. Panorama geral dos dados obtidos no user experience

Através da análise estatística dos dados conseguimos retirar vários resultados para o nosso experimento.

Em relação aos clicks não foi feito nenhum teste estatístico pois descobrimos que o discord tem várias maneiras de chegar às várias funcionalidades do discord por isso o número de clicks acabaria por poder dar uma amostra diferente do que pode ser a realidade, apesar de que em média o número de clicks no género masculino é menor do que o género feminino em qualquer tarefa.

Quando olhamos para os testes Mann-Whitney U test entendemos que a facilidade depende de pessoa para pessoa, tivemos pessoas que fizeram a tarefa num tempo menor em relação a outras, mas introduziram maior dificuldade, o que aponta que a dificuldade depende do senso de cada um.

Por esses fatores achamos que o tempo seja o maior indicador para expressar que existe uma diferença de dificuldade entre os dois grupos, pois se o tempo acaba por ser bastante longo é porque realmente a pessoa não estava a entender como realizar a tarefa.

Olhando para a correlação entre tempo de tarefa e a frequência de utilização, não existiu números expressivos de correlação, isto pode ter acontecido pela amostra ser pequena e também porque para alguns dos nossos usuários do experimento por sentirem que o discord é prático e/ou eficiente acaba por não ser necessário fazer muitas vezes as tarefas para se fazer rapidamente. Apesar disso olhando para os dados um pouco individualmente conseguimos entender que também existe o lado que realmente um utilizador fez poucas vezes uma determinada tarefa e isso vê-se no tempo que o mesmo demorou.

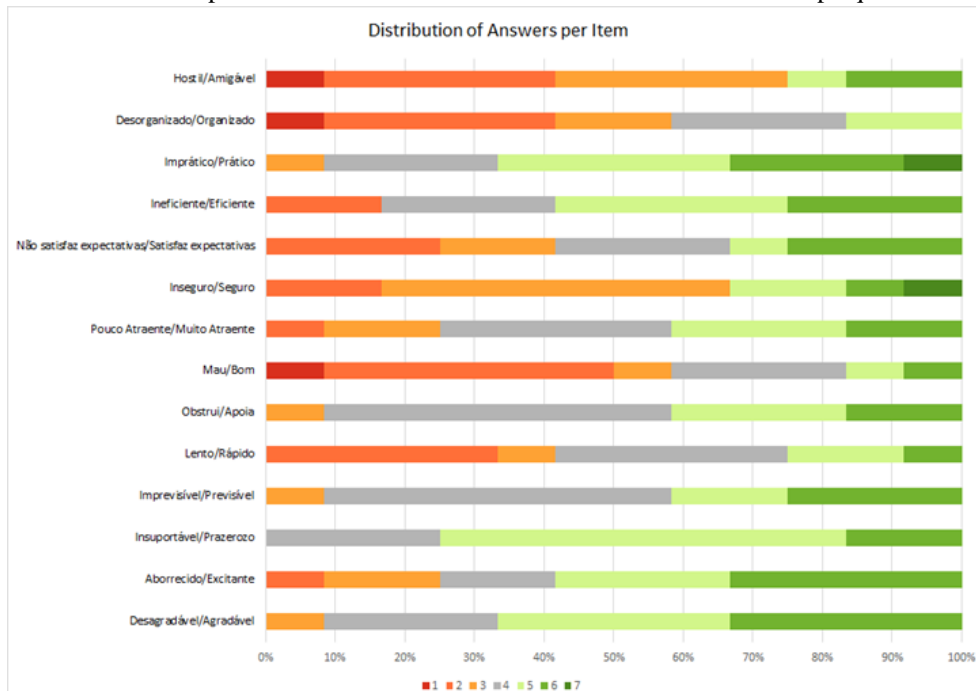


Figura 5. Mostragem dos valores específicos

Em relação ao tempo de conclusão das tarefas dentro dos servidores comunitários segundo gênero, graças ao t-Test, que tem como o p-value (0,004681) e aos dados como média e desvio padrão conseguimos entender que os participantes no estudo que pertencem ao gênero masculino ($\bar{x}=126$; $\sigma=33,6$) têm, em média, tempos totais de conclusão das tarefas 1 a 6 do nosso guião de procedimento do Discord com servidores comunitários menores do que os participantes do gênero feminino ($\bar{x}=243,8$; $\sigma=72,3$). Estes valores também podem de certa forma ligar-se com os resultados das entrevistas e dos questionários, pelos servidores comunitários na sua grande abrangência serem de gaming e existe como sabemos uma baixa percentagem de pessoas do gênero feminino neste mundo, a sua familiaridade com todas as ferramentas dos servidores comunitários pode não ser completa. Como já referido anteriormente se olharmos para a taxa de frequência individualmente podemos ver que acontece em alguns casos nos utilizadores de gênero feminino, não existir frequência de utilização de certas tarefas e de certa forma esses dados acabam por se ligar, isto é, por não haver uma maior aderência a vários servidores, pois os temas que os mesmos contêm são fora dos temas que a pessoa se insere, certas funcionalidades não são usadas.

Além de entendermos se existe ou não diferença entre gêneros, quisemos também entender se a plataforma de pesquisa de servidores dentro do discord realmente se encontra prática o suficiente, e por isso fizemos a tarefa 7 e 8. Nos resultados foi fácil de entender que existe falta de ferramentas na procura dentro do discord, pois nenhum dos nossos utilizadores que fizeram o experimento conseguiu encontrar o servidor.

Enquanto fora do servidor num site de procura de servidores de discord que contém várias ferramentas, como filtrar por membros, reviews e também de “Sorting” caso queiramos procurar apenas por total de membros, ou pelos servers mais recentes por exemplo, tivemos um resultado já bastante expressivo em relação a procura dentro do discord. Apenas 3 pessoas não conseguiram encontrar o discord nesta ferramenta, e apercebemos nos que todas as que não encontraram nunca tinham usado uma ferramenta parecida fora do discord. Os tempos de conclusão da tarefa 8 acabaram por não ser avaliados pois conseguimos entender enquanto fazíamos o experimento que dependia de como a pessoa usava as features do site para demorar mais ou menos tempo, por exemplo se apenas escrevessem a descrição acabariam por encontrar mais rapidamente o discord do que ir por todas as features de selecionar membros o tópico etc...

Apesar disso a outra parte bastante importante era a comparação entre plataformas de procura para conseguirmos retirar o resultado que a procura dentro dos servidores de discord ainda precisa de algumas funcionalidades e que não ajuda o utilizador a pesquisar da melhor maneira possível, seriam necessárias mais ferramentas como existe nos sites externos de procura de servidores.

CONCLUSIONS

Através dos resultados obtidos concluímos que realmente existe uma diferença no tempo de execução das tarefas principais que existem no discord consoante o gênero, sendo que os gêneros masculinos apresentam tempos totais de conclusão das tarefas com servidores comunitários no discord menores, o que nos permite comprovar a nossa expectativa inicial. Também podemos ver que a experiência de utilização acaba por pecar, pois tivemos dados muito abaixo das expectativas.

Em relação a entender se a frequência com o tempo que demoraram acabou por não se conseguir analisar a 100% devido a problemas já relatados anteriormente, talvez com um maior número de utilizadores ajudaria a melhorar estas conclusões, o mesmo acontece com o Mann-Whitney U test. Finalmente conseguimos provar a nossa outra expectativa, a procura de servidores comunitários dentro do discord ainda tem a falta de funcionalidades que ajudam o utilizador a ter uma pesquisa muito mais acessiva e conclusiva, o que já acontece de certa forma em sites de procura de servidores (procura externa).

ANEXOS

FORMULÁRIOS IRB

Os formulários IRB vão ser entregues em anexo.