

O impacto do design em questionários via celular: relações entre modos de luz, quantidade de perguntas e tempo de resposta

The impact of design in mobile surveys: relationships between light modes, number of questions and response time

Gabriela Trombeta

design, celular,
questionários,
modo escuro,
usabilidade

Este estudo avaliou se o design de telas em modo claro (fundo branco com texto escuro) ou escuro (fundo escuro com texto claro) impacta o número de perguntas e tempo que as pessoas respondem a questionários de pesquisa por celular. O estudo contou com processos de design voltado ao usuário (observação, geração de ideias, prototipagem e testagem). Vinte adultos participaram de testes de usabilidade e 120 receberam um link para responder a um longo questionário por celular, devendo clicar em “desistir” quando ficassem cansados. Apesar da maioria dos participantes relatar preferência pelo modo escuro e a média do número de respostas do modo escuro ter sido maior, os testes estatísticos não indicaram diferenças significativas entre os grupos, sugerindo que o uso de um design claro ou escuro não influenciou o suficiente no número de perguntas que as pessoas respondem e no tempo que demoram para fazer isso.

*design, mobile, survey,
dark mode, usability*

This study assessed whether the design of screens in light mode (white background with dark text) or dark mode (dark background with light text) impacts the number of questions and time that people answer mobile surveys. The study relied on human-centered design processes (observation, generation, prototyping and testing). Twenty adults took part in usability tests and 120 received a link to answer a long survey on their cell phones, and clicked on “give up” when they got tired. Though most participants reported a preference for dark mode, and the mean number of dark mode responses was higher, statistical tests did not indicate significant differences between groups. This data suggests that the use of light or dark designs did not sufficiently influence the number of questions that participants answered and the time they took to do that.

1 Introdução

Ao longo dos anos, o uso da internet para aplicação de questionários tem desempenhado um papel cada vez mais importante para pesquisas científicas. O uso de questionários online (i.e., *web surveys*) para obter informações em pesquisas acontece desde os anos 1990

(De Boni, 2020). Entretanto, com o advento da pandemia COVID-19, os questionários online foram vistos como uma alternativa viável à aplicação presencial de questionários em papel, tornando-se predominantes em pesquisas durante os períodos de isolamento social (Hlatshwako et al., 2021).

Questionários online contam com diversos benefícios, como sua velocidade, abrangência e maior quantidade de participantes (De Boni, 2020). Porém, também possuem limitações relevantes. Segundo Budiu e Nielsen (2015), celulares tendem a ser usados em sessões curtas e em diversas circunstâncias e ambientes, o que pode afetar a quantidade de tempo e atenção dada às pesquisas por celular. De acordo com Blumenberg e Barros (2018), uma das maiores limitações de questionários online são suas menores taxas de resposta. Participantes que usam celulares tendem a levar mais tempo para começarem a responder as perguntas e também deixam de concluir questionários em taxas mais altas do que aqueles que respondem por computador (Couper & Peterson, 2017; Mavletova & Couper, 2015).

Um fator que parece interferir nas taxas de resposta e na experiência dos participantes é a plataforma na qual os *web surveys* são preenchidos (Blumenberg et al., 2019), uma vez que podem ser preenchidos em celulares, computadores, navegadores ou aplicativos. No Brasil, a plataforma predominante parece ser o celular, já que dados do Comitê Gestor da Internet no Brasil (2018) indicam que 97% dos brasileiros com acesso à internet fazem seu uso por celular.

Além de questões referentes à plataforma, outros elementos de design dos questionários são relevantes para obtenção de melhores resultados. Em uma revisão sistemática de aspectos de design relevantes para questionários respondidos por celular, Antoun et al. (2018) sugerem que pesquisadores façam um design inicial pensando a menor tela disponível, que o design seja responsivo e apresente-se na vertical. Os autores também recomendam que as perguntas sejam simples, que tarefas complexas sejam simplificadas e questionários em escala (i.e., *grid*) sejam transformados em perguntas de múltipla escolha, de forma que não exijam do participante ações muito complexas e tenham todas as opções de resposta acessíveis. Sugere-se também um máximo de 5 a 7 botões por tela, botões maiores que 6 mm de comprimento e largura para permitir cliques precisos, textos com fonte mínima de 13 pontos e peso mínimo de 4,6 mm para permitir boa legibilidade (ver o artigo completo para ver todas sugestões na íntegra).

A recomendação destes elementos tem grande potencial para auxiliar futuros pesquisadores. Entretanto, conforme ressaltado por De Boni (2020), mais avanços metodológicos ainda são necessários para uma aplicação mais eficiente de questionários online e via celular. Nota-se, pela ausência de resultados na literatura, que um elemento de design que ainda não foi estudado, e parece ser relevante para a disposição das pessoas em responder questionários por celular, é o impacto do uso de designs em modo claro (i.e., *Light Mode*, texto escuro apresentado em fundo claro) ou escuro (i.e., *Dark Mode*,

texto claro apresentado em fundo escuro). Tal elemento parece ser importante considerando que, com o aumento da exposição às telas durante a pandemia, o modo escuro tornou-se popular, sendo constantemente recomendado como uma forma de reduzir a fadiga e aumentar a acuidade e conforto visual (Kim et al. 2019).

Apesar de já existirem alguns estudos sobre este assunto (Kim et al., 2019; Xie et al., 2021), questões referentes ao comportamento das pessoas diante dos diferentes modos de luz ainda não estão claras. No estudo de Xie et al. (2021), fatores biológicos (taxa de intermitência e acomodação da pupila) indicaram redução do cansaço visual durante a leitura de telas no modo escuro, mas, quando perguntados sobre suas preferências, os mesmos participantes relataram preferência e menor cansaço visual diante do modo claro. Isso gerou dúvidas a respeito de que condição seria considerada mais adequada: a que atende às preferências subjetivas de cada um ou a seus resultados biológicos. Além destes achados, Buchner & Baumgartner (2009) afirmam que o modo claro está associado a melhor performance visual e múltiplos estudos (Bauer & Cavonius, 1990; Chan & Lee, 2005; Wang, Fang & Chen, 2003) relatam que tais benefícios são úteis especialmente quando o objetivo é a leitura de textos em telas. Em contrapartida, outros pesquisadores (Rempel & Mautiuk, 2011; Kim et al., 2019) encontraram que o modo escuro reduz a fadiga visual e terceiros (Wang & Chen, 2000) não encontraram diferenças significativas ao comparar os dois modos. Essa diversidade de resultados sugere que ainda não há consenso a respeito dos efeitos de cada modo e qual seria mais adequado para cada contexto.

Considerando que o cansaço visual e as preferências dos participantes podem ter implicações em como eles preenchem questionários, é possível que o design de telas em modo claro e escuro influencie as taxas de resposta em questionários de pesquisa. Tal hipótese, porém, precisa ser testada de forma empírica. Neste sentido, este estudo tem como objetivo avaliar se o design de telas em modo claro (fundo branco com texto escuro) ou escuro (fundo escuro com texto claro) impacta o número de perguntas que as pessoas respondem em questionários por celular e o tempo que levam para fazer isso. Para que esta avaliação seja possível, este estudo buscará responder às seguintes perguntas (i.e., objetivos específicos): 1. Há diferença na quantidade de perguntas que as pessoas respondem em questionários quando usam o modo claro em comparação ao modo escuro? 2. Há diferença no tempo que as pessoas levam para responder questionários quando usam o modo claro e o modo escuro? 3. Será que outras variáveis, como questões demográficas, situações de pressa e a própria preferência dos participantes por um dos modos tem relação com a quantidade de respostas e com tempo que os participantes levam para responder os questionários em cada modo?

A investigação destas perguntas será pautada em processos de Design centrado no usuário (ciclo iterativo de observação, geração de ideias, prototipagem e testagem), uma vez que estes têm como

principal objetivo desenvolver produtos garantindo que (1) atendam às necessidades das pessoas, (2) sejam compreensíveis e utilizáveis, (3) realizem as tarefas esperadas e (4) que a experiência de seu uso seja positiva e agradável (Norman, 2013). Tal abordagem foi escolhida por se encaixar nas características esperadas para questionários de pesquisa que agradem pesquisadores e participantes.

2 Método

2.1 Questões éticas

Essa pesquisa obteve a aprovação do Comitê de Ética da Universidade Federal de São Carlos (CEP/UFSCar), sob o parecer de No 5.307.474

2.2 Participantes

O recrutamento dos participantes foi realizado por meio do método de Amostragem por Bola de Neve (Vanalli, 2014). Para valorizar a heterogeneidade da amostra, o único critério de participação foi que as pessoas tivessem mais de 18 anos.

Participaram desta pesquisa 140 adultos. Destes, 20 participaram de testes de usabilidade do protótipo de cada modo (Figura 1). Nesta etapa, cada rodada de testes foi realizada com 5 pessoas para cada protótipo (Grupo 1 e 2), seguindo recomendações de Norman (2013) de testar protótipos em grupos de cinco pessoas, observar os dados obtidos, fazer aprimoramentos e testar novamente.

Os outros 120 adultos foram contactados posteriormente para participar da tarefa de responder ao questionário e questões complementares (Grupo 3 e 4). Os participantes de cada grupo eram instruídos a chamar outras pessoas para participarem enviando a elas

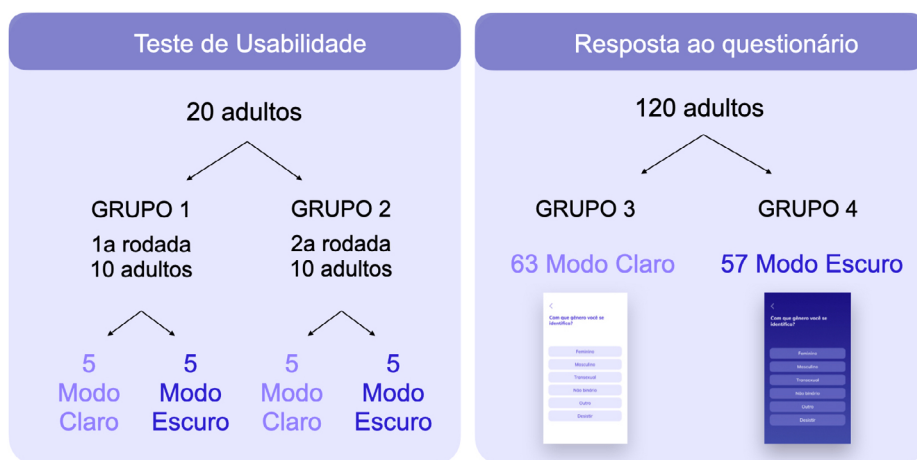


Figura 1 Distribuição de participantes entre grupos e tarefas.

um convite. Este convite continha um link da condição do grupo ao qual pertenciam. Sendo assim, os participantes que responderam ao questionário de modo escuro, por exemplo, chamavam apenas pessoas para participar dessa mesma condição.

2.3 Procedimento

2.3.1 Design e protótipo do questionário

Dois modelos de protótipos (Figura 2) foram desenvolvidos no software Figma. Seu design seguiu as recomendações identificadas na revisão de Antoun et al. (2018, e.g., design inicial para a menor tela disponível, responsivo e na vertical, adaptação de questionários em escala para perguntas de múltipla escolha, máximo de 7 botões por tela, botões maiores que 6 mm de comprimento e largura para permitir cliques precisos, textos com fonte mínima de 13 pontos e peso mínimo de 4,6mm para permitir boa legibilidade) e de Norman (2006, e.g., reduzir instruções ou dá-las apenas uma única vez, manter a tarefa a mesma e oferecer auxiliares mnemônicos para os usuários se lembrarem do que fazer; usar tecnologia para tornar visível o que está acontecendo; possibilidade de desfazer a ação realizada; seguir padronizações já existentes, como ícones em formatos popularmente conhecidos por serem botões, setas conhecidas por voltarem para a página anterior) a fim de aumentar as chances de funcionamento do protótipo inicial.



Figura 2 Exemplo de Tela do questionário do Modo Claro (Esquerda) e Modo Escuro (Direita)

Os elementos visuais foram padronizados para restringir o número de variáveis que poderiam afetar o responder dos participantes, permitindo que apenas as diferenças do modo claro e escuro fossem ressaltadas. Logo, as duas telas eram idênticas, exceto por suas cores, brilho e luminosidade dos elementos. O design destes elementos foi feito seguindo as recomendações da The Web Content Accessibility Guidelines (WCAG) 2.0 (W3C World Wide Web Consortium Recommendation, 2021) e do Google Material Design (Google Inc.). A escolha pelo Google Material Design, em detrimento de outros sistemas, foi informada por dados de que a maior parte da população brasileira está familiarizada com este sistema operacional (Brain & Company, 2019). A acessibilidade e adaptação dos esquemas de cor foram testadas usando testes de cor, contraste e acessibilidade do Adobe Paleta de cores (Anexo 1).

Cada pergunta do questionário foi colocada em uma tela. Todas as telas eram compostas de uma pergunta em destaque, seguida de sete botões com respostas de múltipla escolha. Para cada questão a pessoa podia escolher apenas uma resposta. Ao apertar na opção desejada, ela era levada para próxima pergunta, sinalizando sua interação com aquele botão. Caso desejasse, era possível voltar à pergunta anterior utilizando uma seta presente na lateral esquerda. A possibilidade de desistir da tarefa foi contemplada pelo último botão entre os 7 disponíveis, contendo o texto “Desistir”. Optou-se por fazer com que este botão fosse igual aos outros, para que um design diferente não influenciasse em uma maior ou menor probabilidade de escolha por ele. Instruções sobre o uso destas telas foram dadas apenas uma vez, considerando sugestões de Norman (2006). A opção de exibir uma barra de progresso foi considerada ao longo do estudo, porém, esta foi descartada por ser outro elemento de design que poderia influenciar o comportamento dos participantes, prejudicando a avaliação do efeito dos modos de luz.

2.3.2 *Testagem, observação e geração de ideias*

Para verificar se os protótipos estavam funcionando, foram realizados testes no software Maze (<https://maze.co/>), o qual foi escolhido por suas métricas de usabilidade, tempo de uso de cada tela, mapas de calor e possibilidade de criar diferentes tipos de missão. A escolha por estas métricas baseou-se nas recomendações de Ben Shneiderman (2016) a respeito dos cinco fatores relevantes para avaliar a qualidade de um sistema (tempo de aprendizagem, velocidade: à quantidade e tipo de erros que os utilizadores cometem na utilização do produto; capacidade de manutenção dos conhecimentos; satisfação subjetiva)

Ao realizar o teste no Maze, os participantes dos grupos 1 e 2 deveriam completar três missões referentes às principais funções disponíveis nas telas, sendo elas: responder uma pergunta, voltar a pergunta anterior e desistir de responder (Figura 3). Ao final do teste

Fluxo esperado no teste de usabilidade (Grupo 1 e 2)

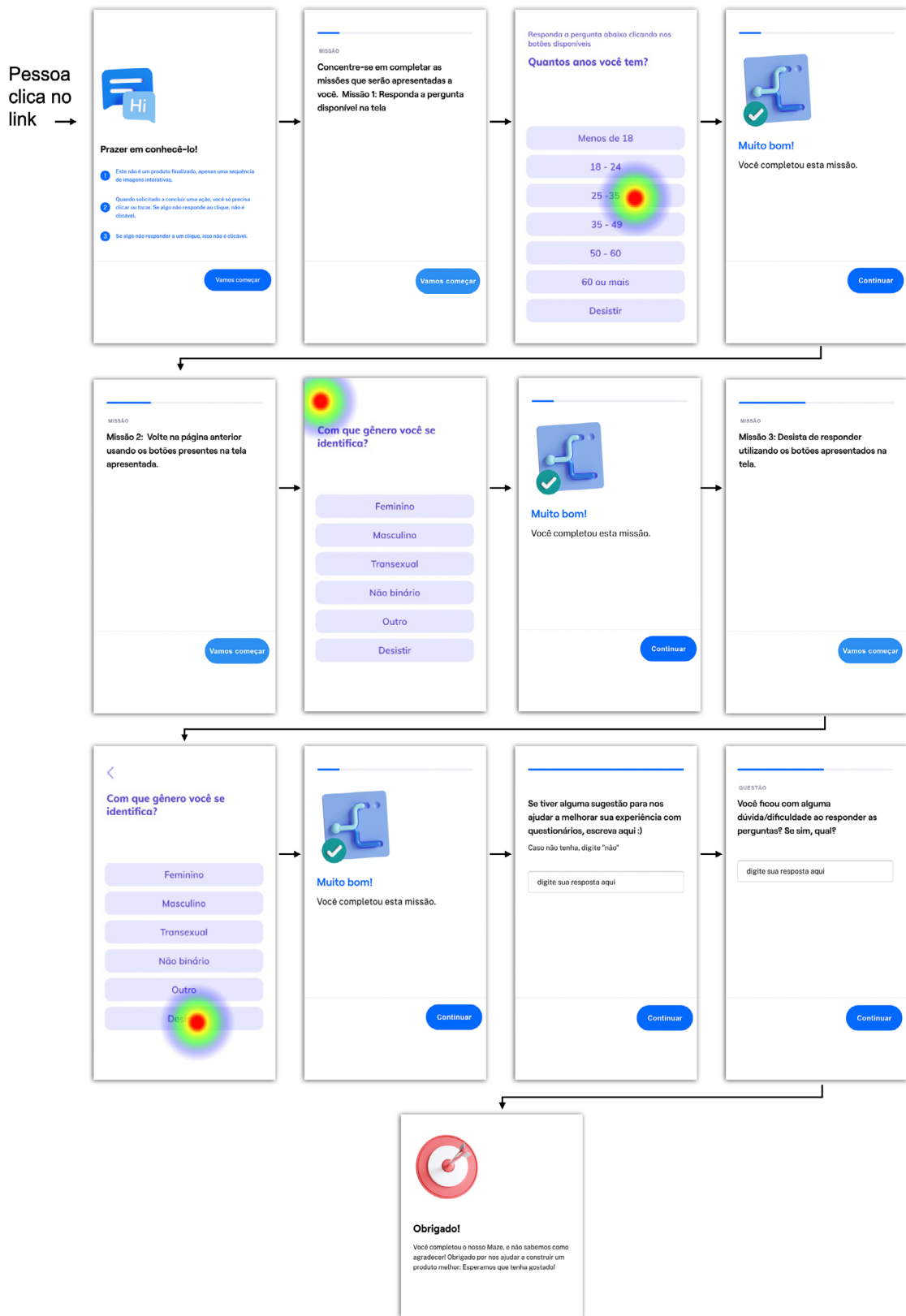


Figura 3 Fluxo de telas do teste de usabilidade do Grupo 1. O Grupo 2 realizou as mesmas atividades porém com a versão das telas do modo escuro.

os participantes também deveriam responder a pergunta: Você teve alguma dúvida/dificuldade em responder o questionário? Se sim, qual?

Quando os resultados do teste de usabilidade indicaram que dois protótipos se demonstraram compreensíveis, um novo teste foi criado dentro do Maze, desta vez com a função de coletar dados que respondessem às perguntas de pesquisa deste estudo.

2.3.3 Coleta de dados

Considerando que o relato das pessoas nem sempre condiz com seus comportamentos, o principal método utilizado para a coleta de dados foi a observação não participante focada em quantas perguntas e por quanto tempo os participantes respondiam ao questionário. Neste método o pesquisador “mantém distância dos eventos observados a fim de evitar influenciá-los” (Flick, 2004, p. 148). Tal procedimento foi considerado como adequado para este estudo considerando que a presença de alguém acompanhando a pessoa no preenchimento do questionário poderia fazer com que ela se sentisse pressionada a responder até o final, e tal situação também não replicaria a condição real no qual a maior parte dos questionários online são respondidos (a distância).

Para que a observação não participante nestas condições fosse possível, um link com o teste contendo o protótipo de cada modo (Figura 4) foi enviado às pessoas de seu grupo. Ao abrir o link, cada pessoa recebia instruções a respeito do que deveria fazer, o objetivo do estudo, a previsão de duração do teste, a afirmação de que os dados coletados seriam anônimos e de que ao clicar em “Vamos participar” estava concordando que leu o Termo de Consentimento (indicado por um link clicável) e estava de acordo em participar da pesquisa.

Os questionários das condições de modo claro e modo escuro continham exatamente as mesmas perguntas, apresentadas na mesma ordem. Como as perguntas presentes nos questionários seriam meramente ilustrativas de perguntas realizadas em pesquisas, este foi composto de perguntas de um outro estudo desenvolvido pela autora. O questionário contemplou um total de 60 questões (Anexo 2), sendo 12 demográficas e 48 questões de instrumentos psicológicos de autorrelato sobre equilíbrio entre demandas do trabalho e da vida pessoal (Escala de Conflito Trabalho-Família, Escala de Manejo de Limites, Escala de Experiências de Descanso) e atenção plena (Inventário Freiburg de Mindfulness). Esse número total de questões foi escolhido por ser considerado grande o suficiente para que houvesse o efeito de cansaço e desistência.

Ao finalizar o questionário também era perguntado ao participante: (1) quanta pressa ele estava sentindo (Escala Likert de 0 a 5, sendo 0 nenhuma pressa e 5 muita pressa), (2) sua preferência por telas no modo claro e no modo escuro (Escala Likert com Emojis) e (3) sua experiência com a pesquisa (dificuldades enfrentadas e sugestões, por

Fluxo de atividades do Grupo 3 e 4

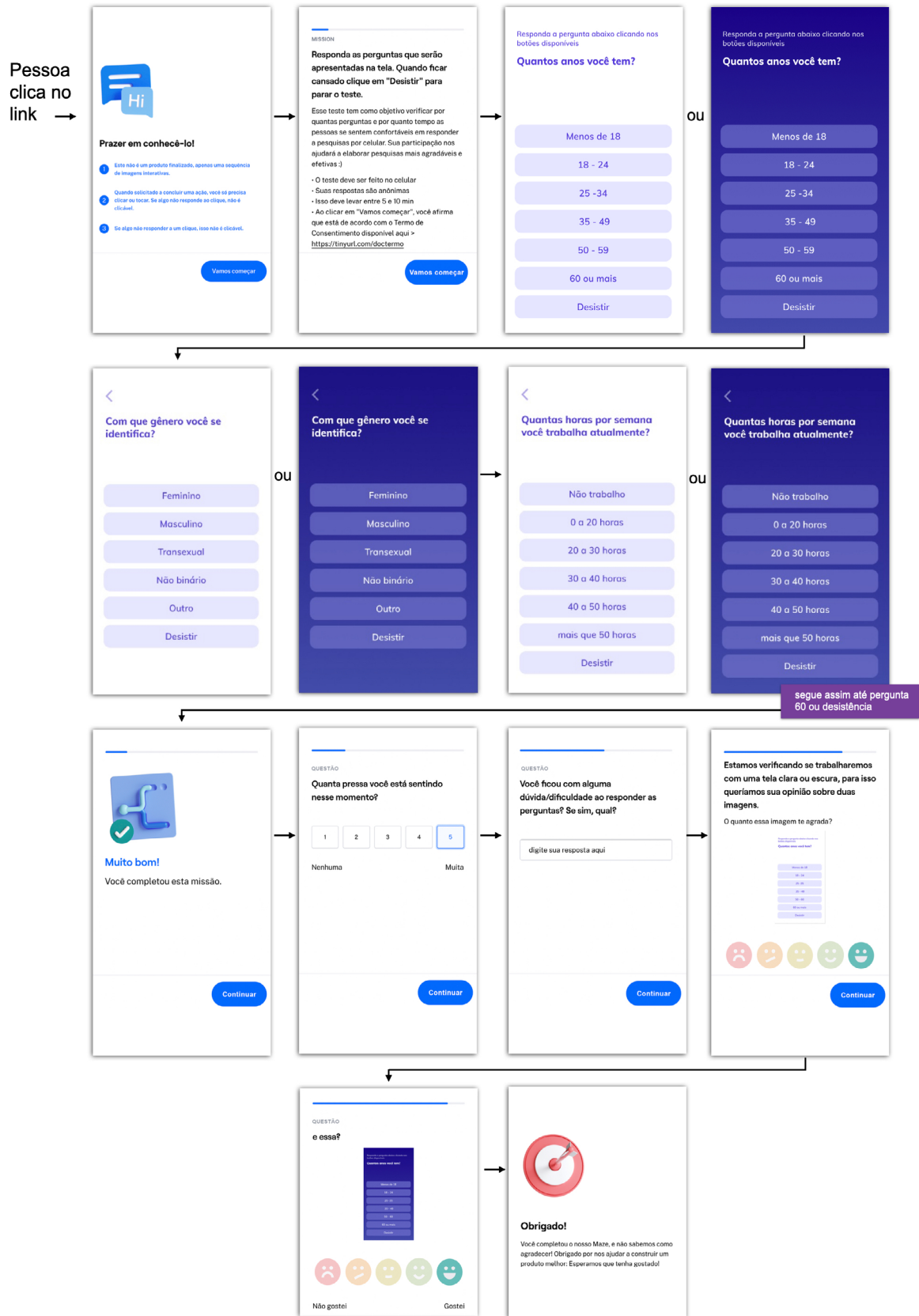


Figura 4 Fluxo de atividades realizadas pelos participantes do Grupo 3 e 4.

meio de respostas abertas). Finalizadas essas questões, o participante recebia uma mensagem de agradecimento por sua colaboração.

2.4 Análise de dados

A análise de dados foi conduzida com os dados coletados pelo Maze.

2.4.1 Testes de usabilidade

Para os teste de usabilidade foram analisadas as medidas de: (1) sucesso direto (porcentagem de usuários que completaram a missão dentro do caminho esperado), (2) indireto (porcentagem de pessoas que completaram a missão por outros caminhos), (3) número cliques errados em cada tela, (4) índices de desistência. No caso de telas com mais de 1% de cliques errados, foram analisados seus mapas de calor, buscando verificar onde o usuário estava tentando clicar. O teste foi considerado bem sucedido quando os resultados atingiram mais de 75% de sucesso direto ou indireto nas missões cumpridas, e menos de 25% de desistências.

2.4.2 Coleta de dados

O número de perguntas respondidas e o tempo de resposta dos participantes de cada grupo foi analisado por meio de dados coletados pelo Maze e tabulados em um arquivo no software SPSS.

A fim de verificar outras variáveis que poderiam afetar o responder dos participantes, os dados demográficos (gênero, idade, se trabalha ou não) e perguntas adicionais (pressa e preferência dos participantes) também foram tabulados e categorizados em números. A presença ou não de trabalho foi tabulada por ser considerada uma variável de identificação dos participantes com o conteúdo do questionário, que era predominantemente voltado a questões de trabalho e vida pessoal. Dados de preferência foram codificados um em relação ao outro, por exemplo, se um participante indicava que gostava muito do modo escuro (5 na escala Likert desse modo), e gostava um pouco do modo claro (4 na escala Likert desse modo), este dado era registrado como “preferência pelo modo escuro”. Quando os números das escalas eram iguais, o dado foi codificado como “sem preferência/igual”. Os dados referentes à experiência dos participantes foram registrados qualitativamente e organizados em tabelas (Anexo 3)

Em prol de responder às perguntas de pesquisa 1 e 2 deste estudo, foram realizadas análises de frequência e testes inferenciais do número de perguntas respondidas e tempo de resposta de cada grupo. Para realização de comparações estatísticas foi verificada a normalidade das amostras e, a partir de seus resultados, foi realizado o Teste U de

Mann-Whitney para dados não paramétricos de grupos independentes. Para responder a pergunta 3, foram realizadas correlações *Spearman* a fim de verificar como as variáveis sexo, idade, pressa e preferência se relacionavam ao número de perguntas e tempo de resposta de cada grupo.

3 Resultados e discussão

3.1 Testes de usabilidade

Foram necessárias duas rodadas de teste de usabilidade para atingir uma experiência desejável. As principais alterações entre a primeira e a segunda rodada foram referentes às instruções dadas e as possibilidades de interação com o protótipo.

Na primeira rodada, a missão 1 (responder uma pergunta) foi concluída com 100% de sucesso direto e 0% de sucesso indireto, desistências e cliques errados. A missão 2 (voltar à pergunta anterior) apresentou 50% de sucesso direto, 0 % de indireto e cliques errados, e 50% de desistências. Já a missão 3 (desistir de responder), contou com 50% de sucesso direto, 25% de indireto, 25% de desistências e 17 cliques errados. Por estes dados e pelos mapas de calor, notou-se que os participantes confundiam as missões a serem cumpridas com objetivo da pesquisa principal (preencher os questionários até onde quisessem). Também aconteceu dos participantes utilizarem os botões do próprio menu do celular para voltar à página anterior, o que não era registrado pelo Maze e gerava problemas na hora de cumprir as missões.

Para corrigir tais erros, o texto de apresentação foi substituído por “concentre-se em cumprir as missões que serão apresentadas à você” e a informação de “Volte a página anterior” foi substituída por “Volte à página anterior usando os botões presentes na tela apresentada”. Tais alterações resultaram em 100% de sucesso direto para a missão 1 e 2 e 80% de sucesso direto e 20 % indireto na missão 3, indicando que a usabilidade desejada foi obtida.

Quanto às dificuldades dos participantes, apesar de 90% dos participantes indicarem que não tiveram dificuldades e que as instruções estavam claras, alguns ainda cometeram erros, o que parece ter ocorrido por conta de falta de atenção na hora de fazer o teste, conforme relatado por uma participante “não prestei muita atenção no comando das missões, achei que precisava só responder diretamente às perguntas que estavam na tela”. Por um lado, apesar de tais resultados dificultarem a eficácia deste teste, por outro, indicaram que o design das telas do questionário estava funcionando no sentido de induzir as pessoas a responderem as perguntas, uma vez que em ambas as rodadas do teste houveram participantes que, apesar de não terem recebido instruções sobre responder às questões, o fizeram da mesma forma.

Tabela 1 Caracterização da amostra (n = 80)

Variáveis	Níveis	Total F.A.	Total F.R. (%)	M. Claro F.A.	M. Claro F.R. (%)	M. Escuro F.A.	M. Escuro F.R. (%)
Gênero	Masculino	52	43,3%	35	55,6%	17	29,8%
	Feminino	64	53,3%	28	44,4%	26	63,2%
	Transexual	1	0,8%	0	0%	1	1,8%
	Não binário	3	2,5%	0	0%	3	5,3%
Idade	18 a 24	25	20,8%	13	21,0%	12	21,1%
	25 a 34	46	38,3%	16	25,8%	30	52,6%
	35 a 49	17	14,2%	13	21,0%	4	7,0%
	50 a 59	19	15,8%	11	17,7%	8	14%
	60 ou +	12	10,0%	9	14,5%	3	5,3%
Trabalho	Não	41	34,2%	23	36,5%	18	31,6%
	Sim	79	65,8%	40	63,5%	39	68,4%
Preferência de cor	Igual	22	20,8%	14	25,5%	8	15,7%
	Claro	32	30,2%	21	38,2%	11	21,6%
	Escuro	52	49,1%	20	36,4%	32	62,7%

Nota: F. A. = frequência absoluta; F. R. (%): frequência relativa; M. Claro = grupo do Modo Claro; M. Escuro = grupo do Modo Escuro.

3.2 Coleta de dados

3.2.1 Caracterização dos participantes

Uma caracterização detalhada da amostra encontra-se descrita na Tabela 1. Pela tabela é possível notar que a amostra foi heterogênea em relação ao gênero, faixa etária e presença ou não de trabalho remunerado nos dois grupos.

3.2.2 Preferência entre os modos

Ao analisar os resultados de cada grupo (Tabela 1), nota-se que os participantes do modo claro preferiram mais o design claro (38,2%) que os participantes do modo escuro (21,6%). Os participantes do modo escuro, por sua vez, preferiram, predominantemente, o design escuro (62,7%). Apesar de ser possível dizer que os participantes pareceram indicar que preferiam o modo ao qual haviam sido expostos, a pequena diferença de preferência entre os participantes do modo claro (21 preferiram o claro vs. 20 preferiram o escuro) faz com que essa afirmação não possa ser feita de forma tão concreta, de forma que seria interessante que estudos futuros investiguem este efeito em amostras maiores.

Ao analisar os dados gerais (soma dos dois grupos), foi registrada uma preferência maior pelo modo escuro (49,1%) em comparação ao claro (30,2%), o que foi um resultado interessante por contrastar com resultados do estudo de Xie et al. (2021), no qual participantes relataram maior preferência por modos claros. Tais diferenças podem ter acontecido pelo presente estudo ter usado modelos em cores, ao invés de modelos em preto e branco como os de Xie et. al. (2021).

3.2.3 Tempo de resposta

A curva de distribuição, o desvio padrão e a média do tempo de resposta entre os grupos (Figura 5) foi bastante semelhante (Claro $M = 387$ s; Escuro $M = 389$ s). Este dado indica que os benefícios do modo claro para a leitura de textos em tela, sugeridos por Bauer e Cavonius (1990), Chan e Lee (2005) e Wang, Fang e Chen (2003), não parecem ter sido relevantes para o tempo levado pelos participantes para completar a tarefa, uma vez que as médias diferiram apenas em dois segundos e os testes comparativos (Tabela 2) não demonstraram diferenças estatisticamente significativas entre os grupos (para ser estatisticamente significativa, o p precisa ser menor que 0,05 e, no caso, para comparação de tempo, o p foi igual a 0,923).

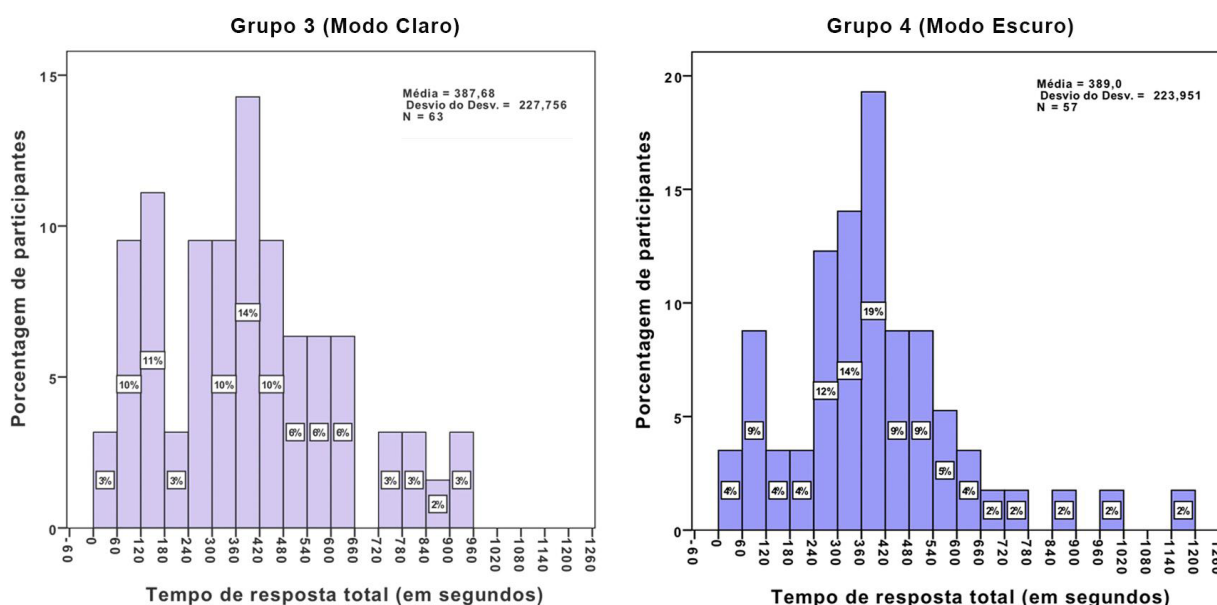


Figura 5 Histograma de porcentagem do tempo de resposta de cada grupo.

Tabela 2 Dados descritivos e inferenciais das análises comparando o tempo de resposta e número de perguntas respondidas entre os grupos

Análises		Tempo de resposta (em segundos)				Número de perguntas			
		M	D.P.	U	p	M	D.P.	U	p
Diferenças entre grupos	Modo Claro	387	227,75	1777	0,923	45,48	19,240	1499	0,080
	Modo Escuro	389	223,95			50,51	17,625		

Nota: M = média; D.P. = desvio Padrão; U = U de Mann Whitney; p = nível de significância estatística; * = $p < 0,05$, resultado estatisticamente significativo; ** = $p < 0,01$, resultado estatisticamente muito significativo.

3.2.4 Número de perguntas respondidas

O grupo que preencheu o questionário no modo escuro obteve uma média maior de perguntas respondidas ($M = 50,51$) do que o modo claro ($M = 45,48$). Também é possível observar que, no modo claro, picos de desistência ocorreram entre as perguntas 10 e 25, e entre as perguntas 55 e 59, já no modo escuro não foi possível identificar picos de desistência, pois estes foram bem distribuídos ao longo das perguntas (Figura 6).

Apesar das diferenças apresentadas acima, os testes estatísticos (Tabela 2) também não identificaram diferenças significativas no número de respostas entre os grupos, sugerindo que o uso de um design claro ou escuro não influenciou o suficiente as pessoas a ponto de gerar diferenças estatisticamente significativas no número de perguntas que elas respondem.

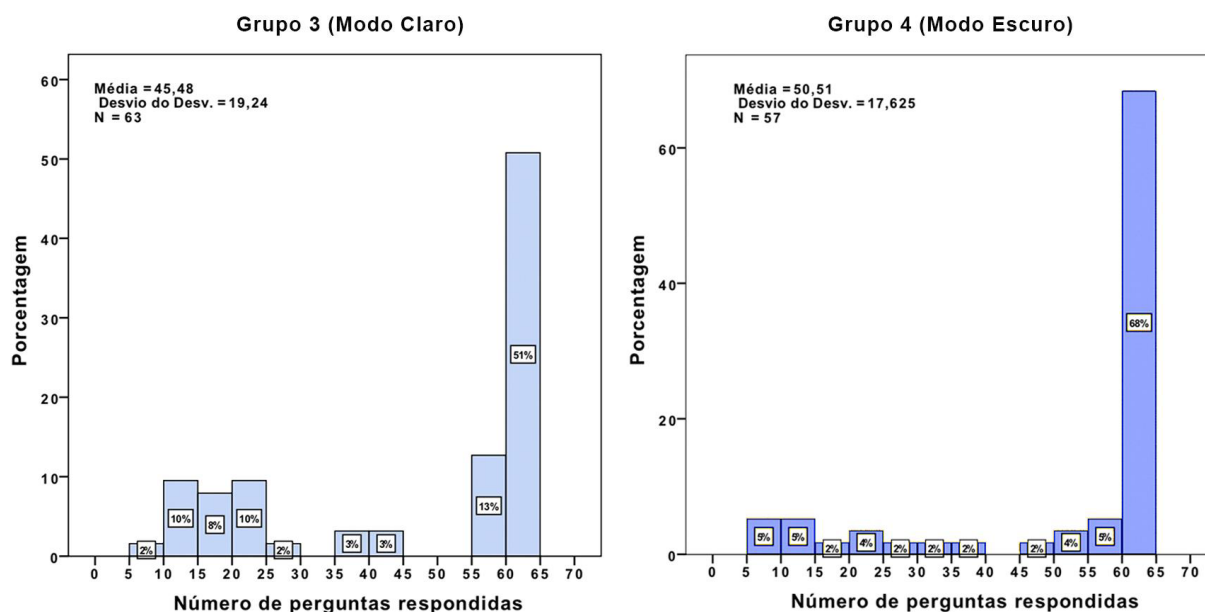


Figura 6 Histograma de porcentagem de perguntas respondidas de cada grupo.

Um aspecto relevante para este dado é que 51% dos participantes do modo claro e 68% dos participantes do modo escuro responderam à todas as perguntas (Figura 6), o que sugere que a ocorrência do efeito teto (i.e., quando a distribuição dos resultados é assimétrica e determinada pela percentagem da população que pontua nos mais elevados níveis das medidas, dificultando a detecção de diferenças entre os grupos). Para evitar que este efeito ocorra em pesquisas futuras, sugere-se que pesquisadores investiguem o comportamento de responder a questionários usando mais de 60 perguntas. Tal recomendação é importante para verificar se haveria diferenças no tempo de resposta e no número de perguntas que as pessoas responderiam em cada modo, se o questionário fosse maior. Apesar de pesquisadores frequentemente considerarem que questionários maiores vão cansar os participantes, estudos internacionais têm demonstrado que participantes respondem a mais perguntas do que é inicialmente esperado (Revilla & Ochoa, 2017; Blumenberg et al., 2019). Estudos como estes ainda são escassos na literatura brasileira, desta forma, entende-se que testes com um maior número de perguntas são importantes para que seja possível determinar com mais clareza as características ideais de design de questionários para população brasileira. Tais dados são extremamente relevantes para que futuras pesquisas possam tomar decisões quanto à duração e comprimento de seus questionários com base em evidências científicas.

3.2.5 Correlações

Não houveram correlações significativas relacionadas à pressa dos participantes, gênero e preferência entre os modos de cores, sugerindo que tais variáveis não estiveram associadas aos comportamentos dos participantes. Por outro lado, foram registradas correlações significativas positivas (Tabela 3) entre as variáveis tempo de resposta e idade (quanto maior a idade maior tempo dedicado a responder as perguntas, $r = 0,31$, $p = 0,001$), entre tempo de resposta e número de perguntas respondidas (quanto mais perguntas respondidas, maior o tempo para responder, $r = 0,64$, $p < 0,001$) e também entre o número de perguntas respondidas e a presença de trabalho, indicando que quando

Tabela 3 Correlações estatisticamente significativas entre as variáveis

	Variáveis					
	Número perguntas	Idade	Gênero	Trabalho	Pressa	Preferência
Tempo de resposta	,64**	,31**				
Número perguntas	,64**			,33**		

Nota: ** = $p < 0,01$, resultado estatisticamente muito significativo.

a pessoa trabalhava ela respondia a mais perguntas ($r = 0,33, p < 0,001$), o que sugere haver uma relação entre a identificação das pessoas com o conteúdo das perguntas e sua disposição em continuar respondendo.

Tais resultados implicam que pesquisas que trabalham com design se beneficiariam de levar em conta tais fatores durante sua coleta de dados, trabalhando, por exemplo, com amostras contemplando faixas etárias específicas a qual seu produto se destina, e se atentando a características demográficas (e.g., se a pessoa trabalha ou não) que possivelmente indiquem identificação ou não da população com o produto/conteúdo que está sendo estudado (e.g., questionários com pergunta sobre o trabalho), já que ambas as variáveis demonstraram-se relevantes para o comportamento dos participantes neste estudo.

3.2.6 *Dificuldades e sugestões dos participantes*

As dificuldades relatadas pelos participantes relacionadas ao design visual do questionário incluíram (1) a ausência de animações que mostrassem a resposta que haviam dado, (2) queixas quanto ao contraste da tela clara, (4) ter que voltar às questões anteriores porque confundiram botões (5), o posicionamento do botão “desistir” abaixo da opção “concordo totalmente” foi mencionado como facilitando a ação de participantes desistirem antes do desejado. As sugestões de design incluíram (1) a presença de animações para sinalizar os botões clicados, (2) a adição de um botão que indicasse incerteza (“não sei dizer”), (3) e a adição de uma barra de progresso. Uma lista completa com as dificuldades e sugestões dos participantes (incluindo questões relacionadas ao conteúdo das perguntas), encontra-se disponível no Anexo 3. Sugere-se que pesquisadores que trabalham com questionários considerem incorporar estes elementos no design de suas pesquisas.

3.3 Limitações

Apesar do número de participantes entre os grupos ter sido similar e a população ter sido heterogênea em ambos, sugere-se que estudos futuros controlem sistematicamente as características demográficas de cada grupo. Reconhece-se também que o uso de recrutamento de amostragem por bola de neve conta com questões de viés relacionados a aceitabilidade social. Neste sentido, a fim de que seja possível verificar a generalização dos resultados encontrados, recomenda-se que estudos futuros explorem outros tipos de recrutamento de participantes, verificando se os resultados desta pesquisa poderiam ser replicados se o estudo fosse conduzido utilizando outros tipos de delineamento. Sugere-se também que sejam feitas novas investigações sobre os possíveis efeitos de modos claros e escuros na desistência dos participantes caso a quantidade de perguntas fosse maior do que 60. Por fim, também torna-se importante ressaltar que os resultados

desta pesquisa são referentes a um questionário contendo questões demográficas e psicológicas de autorrelato, com opções de múltipla escolha, sendo necessários mais estudos para verificar se os dados encontrados seriam generalizáveis para questionários contendo perguntas de outros tipos.

4 Considerações finais

Este estudo teve como objetivo avaliar se o design de telas em modo claro (fundo branco com texto escuro) ou escuro (fundo escuro com texto claro) impacta de formas diferentes o número de perguntas e tempo que as pessoas respondem a questionários de pesquisa por celular. Tal avaliação foi feita buscando responder a três perguntas de pesquisa: 1. Há diferença na quantidade de perguntas que as pessoas respondem em questionários quando usam o modo claro em comparação ao modo escuro? 2. Há diferença no tempo que as pessoas levam para responder questionários quando usam o modo claro e o modo escuro? 3. Será que outras variáveis, como questões demográficas, situações de pressa e a própria preferência dos participantes por um dos modos, tem relação com a quantidade de respostas e com tempo que os participantes levam para responder os questionários em cada modo?

Em resposta à pergunta 1 e 2, os resultados deste estudo indicam que, apesar da maioria dos participantes relatar preferência pelo modo escuro e a média do número de respostas do modo escuro ter sido maior, não foram identificadas diferenças estatisticamente significativas entre os grupos. Isso significa que não houve diferença na quantidade de perguntas que as pessoas responderam, nem no seu tempo de resposta, de acordo com o modo de luz utilizado. Esse resultado sugere, portanto, que o design de modo claro e escuro tiveram efeitos semelhantes no comportamento das pessoas.

Na prática, entende-se que designers podem olhar para estes dados e interpretá-los de duas maneiras. A primeira maneira seria valorizar os resultados estatísticos obtidos como a evidência científica que temos disponível no momento, o que significa considerar a informação de que “não houve diferenças estatísticas entre os modos de luz” e portanto, ao trabalhar no design de questionários online seria possível optar por qualquer um deles, considerando que seus efeitos serão semelhantes. A segunda possibilidade seria levar a primeira opção em consideração, mas também valorizar o dado de que, em geral, os participantes relataram maior preferência pelo modo escuro e, apesar deste dado não ser estatisticamente significativo, é possível que se tenha mais chances de agradar as pessoas priorizando um design que contemple este modo. Também é necessário considerar que embora esta pesquisa não tenha encontrado diferenças significativas entre o uso dos dois modos, ela é a primeira pesquisa sobre o assunto realizada na literatura brasileira, sendo necessário mais estudos para que seus dados possam

ser generalizados com mais confiança. Neste sentido, ressalta-se a necessidade de profissionais buscarem estar constantemente atualizados a respeito de estudos sobre o assunto, pois novas pesquisas podem surgir e apresentar resultados que reforçam estes dados, ou que os contradizem, gerando diferentes implicações práticas.

Quanto à pergunta 3, observou-se que os participantes que trabalhavam responderam mais perguntas que os participantes que não trabalhavam (i.e., correlação positiva entre trabalho e número de perguntas respondidas). Considerando que o conteúdo do questionário referia-se ao trabalho, entende-se, portanto, que as pessoas que se identificaram com o conteúdo das questões responderam a mais perguntas que as pessoas que não se identificavam com o tema que estava sendo abordado. Na prática, isso significa que é aconselhado que profissionais que buscam realizar pesquisas criem mecanismos para verificar se a população que está sendo investigada se identifica com o tema que está sendo pesquisado, por conta de que essa ausência de identificação pode prejudicar o engajamento, gerando desistências para continuar a responder. Além deste fator, também recomenda-se que profissionais que trabalham com questionários reconheçam o possível viés dos dados coletados, uma vez que apenas a população que se identifica com o conteúdo dos questionários tende a respondê-los até o final.

Como as outras variáveis investigadas (pressa dos participantes, gênero e preferência entre os modos de cores) não afetaram significativamente o número e o tempo das respostas dos participantes, sugere-se que novas pesquisas investiguem outras variáveis que podem interferir no comportamento de responder questionários online, tais como outras questões demográficas (como nível de escolaridade, por exemplo) e questões referentes ao conteúdo do questionário. Em termos de design visual, sugere-se que novos estudos busquem testar empiricamente se as sugestões dadas pelos participantes (animação dos botões, barras de progresso) podem melhorar suas experiências com questionários. Pesquisas a respeito destas novas variáveis e elementos de design são importantes para que seja possível detectarmos cientificamente pontos aos quais pesquisadores e profissionais precisam se atentar no design de questionários online, em prol de facilitar a participação das pessoas, melhorar seus níveis de participação e aumentar suas taxas de resposta.

Referências

- Antoun, C., Katz, J., Argueta, J., & Wang, L. (2018). Design heuristics for effective smartphone questionnaires. *Social Science Computer Review*, 36(5), 557–574. <https://doi.org/10.1177/0894439317727072>
- Bauer, D., & Cavonius, C. R. (1990). *Ergonomic aspects of visual display terminals*. London: Taylor & Francis, Chapter: Improving the legibility of visual display units through contrast reversal.

- Blumenberg, C., & Barros, A. (2018). Response rate differences between web and alternative data collection methods for public health research: A systematic review of the literature. *International Journal of Public Health*, 63(6), 765–773. <https://doi.org/10.1007/s00038-018-1108-4>
- Blumenberg, C., Menezes, A. M. B., Gonçalves, H., Assunção, M. C. F., Wehrmeister, F. C., Barros, F. C., & Barros, A. J. D. (2019). The role of questionnaire length and reminders frequency on response rates to a web-based epidemiologic study: A randomised trial. *International Journal of Social Research Methodology*, 22(6), 625–635. <https://doi.org/10.1080/13645579.2019.1629755>
- Brain & Company (2019). *Impacto econômico e social do Android no Brasil*. https://www.bain.com/contentassets/20e0815cfd784b4a8dead63475b42380/v02_b-impactos-do-android-no-brasil.pdf
- Buchner, A., & Baumgartner, N. (2007). Text – background polarity affects performance irrespective of ambient illumination and colour contrast. *Ergonomics*, 50(7), 1036–1063. <https://doi.org/10.1080/00140130701306413>
- Budiu, R., & Nielsen, J. (2015). *User experience for mobile applications and websites* (3rd ed.). Nielsen Norman Group. <http://www.nngroup.com/reports/mobile/>
- Chan, A. H. S., & Lee, P. S. K. (2005). Effect of display factors on Chinese reading times, comprehension scores and preferences. *Behaviour and Information Technology*, 24(2), 81–91. <https://doi.org/10.1080/0144929042000267073>
- Comitê Gestor da Internet no Brasil (2018). Pesquisa sobre o uso das tecnologias de informação e comunicação nos domicílios brasileiros: TIC Domicílios 2018. Disponível em: https://www.cetic.br/media/docs/publicacoes/2/tic_dom_2017_livro_eletronico.pdf
- Couper, M. P., & Peterson, G. J. (2017). Why do web surveys take longer on smartphones? *Social Science Computer Review*, 35(3), 357–377. <https://doi.org/10.1177/0894439316629932>
- De Boni, R. B. (2020). Websurveys nos tempos de COVID-19. *Cadernos de Saúde Pública*, 36(7). <https://doi.org/10.1590/0102-311X00155820>
- Flick, U. (2004). *Uma introdução à pesquisa qualitativa* (2. ed.). Porto Alegre: Bookman.
- Google, Inc. (2021). Google Material Design: Dark Theme. Disponível em: <https://material.io/design/color/dark-theme.html>
- Hlatshwako, T. G., Shah, S. J., Kosana, P., Adebayo, E., Hendriks, J., Larsson, E. C., Hensel, D. J., Erausquin, J. T., Marks, M., Michielsen, K., Saltis, H., Francis, J. M., Wouters, E., & Tucker, J. D. (2021). Online health survey research during COVID-19. *The Lancet. Digital health*, 3(2), e76–e77. [https://doi.org/10.1016/S2589-7500\(21\)00002-9](https://doi.org/10.1016/S2589-7500(21)00002-9)
- Kim, K., Erickson, A., Lambert, A., Bruder, G., & Welch, G. (2019). ‘Effects of dark mode on visual fatigue and acuity in optical see-through head-mounted displays. *Proc. Symp. Spatial User Interact.*, New Orleans, LA, USA, Oct. 2019, pp. 1–9. <https://doi.org/10.1145/3357251.3357584>
- Mavletova, A., & Couper, M. P. (2015a). A meta-analysis of breakoff rates in mobile web surveys. In: D. Toninelli, R. Pinter, & P. de Pedraza (Eds.), *Mobile research methods: Opportunities and challenges of mobile research methodologies* (pp. 81–98). London, England: Ubiquity Press. <http://dx.doi.org/10.5334/bar.f>

- Norman, D. A. (2013). *The design of everyday things: Revised and Expanded Edition*. New York: Basic Books.
- Norman, D. A. (2006). *O Design do dia-a-dia*. Rio de Janeiro: Rocco. Capítulo Sete: Design Centrado no Usuário.
- Rempel, A. G., & Mautiuk, R. (2011). Display considerations for improved night vision performance. In *Proc. 19th Color Imag. Conf.: Color Sci. Eng. Syst., Technol., Appl. (CIC)*, Atlanta, CA, USA, 2011, pp. 191–194. <https://www.cl.cam.ac.uk/~rkm38/pdfs/rempel11dcinvp.pdf>
- Revilla, M., & Ochoa, C. (2017). Ideal and maximum length for a web survey. *International Journal of Market Research*, 59(5), 557–565. <https://doi.org/10.2501/IJMR-2017-039>
- Wang, A. H., & Chen, M. T. (2000). Effects of polarity and luminance contrast on visual performance and VDT display quality. *Int. J. Ind. Ergonom.*, 25(4), 415–421, [https://doi.org/10.1016/S0169-8141\(99\)00040-2](https://doi.org/10.1016/S0169-8141(99)00040-2)
- Wang, A. H., Fang, J. J., & Chen, C. H. (2003). Effects of VDT leading-display design on visual performance of users in handling static and dynamic display information dual-tasks. *International Journal of Industrial Ergonomics*, 32(2), 93–104. [https://doi.org/10.1016/S0169-8141\(03\)00041-6](https://doi.org/10.1016/S0169-8141(03)00041-6)
- W3C World Wide Web Consortium Recommendation (2021). Web Content Accessibility Guidelines (WCAG) 2.0. <https://www.w3.org/TR/AERT/#color-contrast>
- Xie, X., Song, F., Liu, Y., Wang, S., & Yu, D. (2021), Study on the Effects of Display Color Mode and Luminance Contrast on Visual Fatigue. In: *IEEE Access*, vol. 9, pp. 35915–35923. <https://doi.org/10.1109/ACCESS.2021.3061770>

Sobre a autora

Gabriela Trombeta

gabriela_trombeta@hotmail.com

Universidade Federal de São Carlos (UFSCar)

Programa de Pós-Graduação em Psicologia

Artigo recebido em/*Submission date*: 23/3/2022

Artigo aprovado em/*Approval date*: 16/5/2022