## UNIVERSIDADE DE SÃO PAULO

Escola de Comunicações e Artes "Especialização em Pesquisa de Mercado Aplicada em Comunicações".

## **LUCAS CHAVES**

## O IMPACTO DO FORMATO DE APLICAÇÃO DO QUESTIONÁRIO NO RESULTADO

## UNIVERSIDADE DE SÃO PAULO

Escola de Comunicações e Artes "Especialização em Pesquisa de Mercado Aplicada em Comunicações".

## **LUCAS CHAVES**

## O IMPACTO DO FORMATO DE APLICAÇÃO DO QUESTIONÁRIO NO RESULTADO

Monografia apresentada ao Departamento de Relações Públicas, Propaganda e Turismo da Escola de Comunicações e Artes da Universidade de São Paulo, em cumprimento parcial às exigências do Curso de Pós-Graduação-Especialização, para obtenção do título de especialista em "Pesquisa de Mercado Aplicada em Comunicações", sob orientação da profa. Ms. Nelma Soares.

À minha família e colegas de profissão.

## **AGRADECIMENTOS**

À minha esposa pelo apoio e companheirismo nas horas dedicadas.

Aos colegas de curso, pela ajuda e participação.

Aos colegas de trabalho, que contribuíram para a evolução do interesse nesta carreira.

**RESUMO** 

CHAVES, L. O impacto da aplicação de atributos de concordância no resultado obtido.

2012,44 f. Monografía (Especialização) – Escola de Comunicação e Artes, Universidade de

São Paulo, São Paulo, 2012.

O projeto tem o intuito de testar uma hipótese, verificando se há diferença significativa em

diferentes formas de aplicação na avaliação de atributos em questionários estruturados de

pesquisa de mercado quantitativa. A ideia permeia situações onde atributos de concordância e

opinião, são aplicados isoladamente em um universo e na sequencia aplicados em outro

formato, no mesmo universo. Além do fator quantidade de respostas, testar se o impacto da

influência entre os atributos, também mudaria a nota na avaliação final.

Palavra-chave: Questionário. Atributos. Hipótese. Pesquisa. Quantitativa.

**ABSTRACT** 

CHAVES, L. The impact of applying attributes of agreement in the result. 2012.45 f.

Monograph (Specialization) - School of Communication and Arts, University of São Paulo,

São Paulo, 2012.

The project is intended to test a hypothesis by checking whether there are significant

differences in different forms of application in the evaluation of attributes in structured

questionnaires quantitative market research. The idea permeates situations where attributes of

agreement and opinion are applied alone in a universe and applied in the second time to

another format in the same public profile. In addition to the quantity answers factor, to test the

influence between the attributes also change the note on the final evaluation.

**Keywords:** Questionnaire. Attributes. Hypothesis. Research. Quantitative.

## SUMÁRIO

	INTR	ODUÇÃO	12
	MET	ODOLOGIA	14
1	QUES	STIONÁRIO DE CAMPO	16
	1.1	Objetivos do Questionário	17
	1.2	Pesquisa Online	18
	1.3	Métricas	19
	1.4	Escala Likert	21
2	AMO	STRA E COLETA	25
	2.1	Amostra	25
	2.2	Produtividade de campo	26
3	ANÁ	LISE DESCRITIVA	28
	3.1	Questionário 1	28
	3.2	Questionário 2	30
	3.3	Comparabilidade	32
	3.4	Teste de médias	33
	3.5	Teste t de Student	35
4	RESU	JLTADOS	37
	4.1	Atributo 1 – Saúde pública	37
	4.2	Atributo 2 – Segurança pública	39
	4.2	Atributo 3 – Educação pública	42
	CON	CLUSÃO	45
	REFI	ERÊNCIAS	47

## LISTA DE FIGURAS

Figura 1	Modelo do Questionário – www.encuestafacil.com	17
Figura 2	Escala Likert - Fishbein and Ajzen, 1975	23
Figura 3	Questionário com 3 atributos aplicados	26
Figura 4	Questionário com 3 atributos aplicados	27
Figura 5	Gráfico da amostra por sexo – Fase 1	29
Figura 6	Gráfico da amostra por escolaridade – Fase 1	29
Figura 7	Analise descritiva questionário - Fase 1	29
Figura 8	Gráfico da amostra por sexo – Fase 2	31
Figura 9	Gráfico da amostra por escolaridade – Fase 2	31
Figura 10	Analise descritiva questionário - Fase 2	32
Figura 11	Analise descritiva por sexo – Fase 1	33
Figura 12	Analise descritiva por sexo – Fase 2	33
Figura 13	Box Plot – Comparação entre os grupos – Atributo Saúde	37

Figura 14	Box Plot – Comparação entre os grupos – Atributo Segurança	39
Figura 15	Box Plot – Comparação entre os grupos – Atributo Educação	42

## INTRODUÇÃO

Na área de pesquisa mercadológica, em pesquisas quantitativas, muito se discute sobre as melhores formas de abordagem do entrevistado por meio de um questionário estruturado. Que tipo de escala utilizar, que tipo de avaliação aplicar (concordância, satisfação, recomendação, etc.) e a melhor forma de absorver a real percepção do entrevistado.

No mercado consumidor de pesquisas de mercado, existe uma saturação na aplicação de questionários estruturados, em relação ao conteúdo extraído, as escalas utilizadas e principalmente em relação ao volume de informações extraídas.

Hoje em dia, em uma pesquisa de mercado quantitativa, ouvimos alguns paradigmas sobre o tempo de aplicação. Este tempo poderá variar muito de acordo com a abordagem, sendo *Catti* (Telefone), Pessoal, Web, eles permitem um tempo de tolerância do entrevistado muito volátil.

Reter o entrevistado em uma pesquisa por telefone, esta a cada dia mais difícil, sendo que para extrair informações mais concisas em curto espaço de tempo, tornou-se mais complexo. O entrevistado na maioria das vezes acaba seguindo um ritmo de avaliação, se abstendo ou mesmo interrompendo a entrevista e assim, causando um prejuízo às empresas de campo e aos contratantes do estudo, tanto na qualidade das informações quanto no custo operacional.

Outro ponto importante estudado neste projeto é o impacto das questões precedentes em um questionário, na avaliação das questões subsequentes. Um exemplo disto seria questionar ao entrevistado o quanto ele concorda que um Porsche é um carro veloz, de duas formas, uma isolada e outra seguinte à mesma pergunta, avaliando uma Ferrari. Com certeza existiria para alguns dos entrevistados a avaliação sobre comparabilidade.

E por fim, outro aspecto técnico que pode contribuir na avaliação seria a quantidade de perguntas (atributos) avaliada, que em grande volume pode gerar uma fadiga mental no entrevistado e interferir no resultado esperado.

A intenção neste estudo, é avaliar se questões mais objetivas e mais representativas, podem substituir o volume de conteúdo de avaliações, que uma grande bateria pode oferecer em um questionário. Testando a variabilidade entre diferentes formas de aplicação, de mesmas perguntas e contextos distintos. Pensando na premissa estabelecida, isso pode

<sup>&</sup>lt;sup>1</sup> Cati - Computer assisted telephone interviewing

ocasionar grande economia financeira, de tempo e de produtividade nas pesquisas mercadológicas e principalmente, maior precisão das informações coletadas.

#### **METODOLOGIA**

A proposta deste estudo é testar a hipótese presumida, de que existe diferença significativa na avaliação de atributos de qualidade, satisfação, opinião, utilizando dois caminhos diferentes.

Estabelecendo dois distintos caminhos de abordagem, considerando a aplicação em um mesmo universo, com mesma metodologia de coleta em campo (online), aplicaremos 3 atributos idênticos, da seguinte forma:

- 1. Três atributos de tema popular em bateria de 3 componentes
- 2. Três atributos de tema popular em bateria de 17 componentes

A intenção é comprovar que existe diferença nos resultados, utilizando dois formatos de abordagem, onde o mesmo universo de pessoas respondeu o primeiro questionário e após 1 mês o segundo questionário.

O critério para definição dos atributos a serem avaliados, foi o conhecimento público e de interesse social, por se tratar de serviços básicos da população.

O questionário foi aplicado considerando o mesmo universo, onde o perfil do público de campo foi aleatório em usuários de internet, de duas redes sociais específicas. Na coleta tivemos valores distintos de retorno, pois para maior comodidade e discrição do entrevistado, procuramos não exigir a identificação. Desta forma, não afirmaremos que exatamente o mesmo grupo de pessoas respondeu aos dois questionários, mas podemos afirmar que o mesmo universo respondeu aos dois questionários.

O disparo foi realizado por meio da internet, contemplando apenas os usuários das duas redes sociais estipuladas, aproximadamente 1.000 pessoas envolvidas. Por se tratar de uma pesquisa de autopreenchimento, procuramos balancear pelo sexo, quantidade semelhante dentro das duas coletas. As datas de envio foram entre Nov/2011 e Jan/2012.

O objeto de estudo pretende avaliar se o excesso de perguntas pode influenciar na avaliação do entrevistado, por conta disto a abordagem sócio-demográfica se limitou ao levantamento do sexo e do grau de escolaridade. Em virtude de limitação técnica da ferramenta, não foi utilizado o esquema de rodízio entre os atributos pesquisados.

Os atributos testados foram:

- "A saúde pública em nosso país é adequada à nossa necessidade";
- "A segurança pública em nosso país é adequada à nossa necessidade";

"A educação pública em nosso país é adequada à nossa necessidade"

Em paralelo, em um dos grupos testados aplicamos além dos três atributos acima citados, outros 14 atributos, no intuito de testar a influência e comparabilidade no momento da avaliação, além do impacto no entrevistado, do volume de perguntas disponibilizadas.

#### São eles:

- "A previdência social em nosso país é adequada a nossa necessidade";
- "A política social em nosso país é adequada a nossa necessidade";
- "As eleições em nosso país são justas";
- "A saúde em nosso país é uma das melhores do mundo"
- "O investimento em esporte e cultura em nosso país é um dos maiores do mundo"
- "A segurança pública em nosso país é uma das melhores do mundo"
- "A corrupção em nosso país é uma das menores do mundo"
- "A educação pública em nosso país é uma das melhores do mundo"
- "O transporte público no Brasil é adequado à nossa necessidade"
- "As leis tributárias no Brasil são justas"
- "O transito de veículos nas vias públicas é seguro"
- "A presidente Dilma Roussef, está cumprindo o prometido em campanha eleitoral".
- "O desemprego tem diminuído a cada dia no Brasil"
- "A violência tem diminuído a cada dia no Brasil"
- "O salário mínimo no Brasil é um dos mais justos do mundo"
- "A qualidade dos serviços públicos tem melhorado a cada ano no Brasil
- "A mídia em nosso país visa a informação do cidadão

Com os dados coletados, testaremos a seguinte hipótese:

H0: As medias entre os atributos nas duas aplicações são iguais

H1: As medias entre os atributos nas duas aplicações não são iguais

Resumindo, queremos saber se mesmas perguntas realizadas de formas diferentes, mas com mesmo enunciado, geram diferentes resultados.

## 1 QUESTIONÁRIO DE CAMPO

O questionário utilizado em campo é o formulário com as perguntas a serem avaliadas pelo grupo pesquisado. Neste estudo utilizamos dois tipos de questionários que conheceremos na sequencia do capítulo.

Esta é a ferramenta mais importante utilizada neste estudo. Por meio desta ferramenta conseguimos obter as informações necessárias para o teste realizado, que neste caso aconteceu em dois momentos. Temos a seguir, o formato dos questionários utilizados e os temas abordados.

Foram aplicados dois questionários, onde o primeiro contempla apenas os 3 atributos a serem testados e duas perguntas de perfil.

#### Questionário 1:

- Sexo
- Grau de Instrução
- Atributos (3)

#### Questionário 2:

- Sexo
- Grau de Instrução
- Atributos (17)

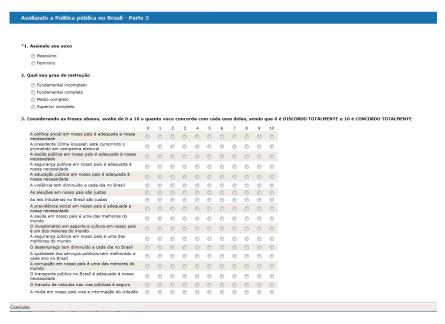
A escala utilizada para os atributos foi a *Likert*, 0 a 10 pontos, com o ponto intermediário (5).

Batizada em homenagem ao seu criador, Rensis Likert, a escala Likert é uma escala amplamente utilizada que exige que os entrevistados indiquem um grau de concordância ou discordância com cada uma da serie de informações sobre o objeto de estimulo.

A ferramenta de envio do questionário foi o site www.encuestafacil.com. Como líder mundial de pesquisa na Internet, a Decision Analyst realizou centenas de surveys internacionais online de mais de 3,5 milhões de consumidores nos Estados Unidos, Canadá, Europa, América Latina e Ásia. Com base em suas próprias experiências e em observações de outros profissionais, a Decision Analyst acredita que as pesquisas pela internet oferecem

dados de melhor qualidade, são mais rápidas e menos onerosas do que as realizadas por telefone.

Figura 1: Modelo do Questionário – www.encuestafacil.com



Fonte: www.encuestafacil.com

Nos tópicos seguintes deste capítulo, conheceremos os objetivos deste questionário, entendendo melhor como ele foi aplicado. Descrevemos também como funciona uma pesquisa online, como é utilizada hoje em dia com suas vantagens e desvantagens, quais são as métricas utilizadas na construção de um questionário de pesquisa de mercado e o estilo de escalas utilizadas.

#### 1.1 Objetivos do Questionário

Qualquer questionário tem 3 objetivos específicos (MALHOTRA, 2004). Em primeiro lugar, deve transformar a informação desejada em um conjunto de perguntas específicas em que os entrevistados tenham conhecimento para responder. A elaboração de perguntas que os entrevistados saibam responder, respondam efetivamente e que tragam a informação necessária, é uma tarefa difícil. Duas maneiras aparentemente semelhantes de formular uma pergunta podem gerar informações divergentes. Assim, este objetivo é sempre um grande desafio.

Em segundo lugar, um questionário precisa motivar e incentivar o entrevistado a se deixar envolver pela entrevista, principalmente em casos de autopreenchimento. Entrevistas incompletas têm, na melhor das hipóteses, uma utilidade precária. Ao planejar um questionário, o entrevistador precisa sempre minimizar o cansaço e o tédio do entrevistado, esforçando-se também para minimizar as respostas incompletas e as não respostas.

Em terceiro lugar, um questionário deve sempre minimizar o erro de resposta, que surge quando os entrevistados dão respostas imprecisas ou quando elas são registradas ou analisadas incorretamente.

Neste trabalho, buscou-se aperfeiçoar a abordagem ao entrevistado, evitando o cansaço desnecessário, que não fosse inerente ao objeto de estudo.

Procurando envolver e motivar o entrevistado, o tema escolhido é de conhecimento público, com abordagem polêmica e de fácil opinião. Os temas escolhidos estão voltados à administração pública brasileira, como saúde, segurança, educação e política, entre outros.

Com estes objetivos, o estudo procurou levantar as informações necessárias para o teste realizado.

## 1.2 Pesquisa Online

As pesquisas on-line são cada vez mais comuns em empresas deste segmento, porém ainda é restrito às classes A, B e em menor escala para a classe C. Este fato se dá por causa da baixa penetração da Internet no Brasil e pela dificuldade das classes sociais de baixa renda em adquirir um computador. Problemas sociais à parte, a pesquisa on-line trouxe mais comodidade e agilidade em relação ao método convencional — entrevistador e entrevistado ou ligações telefônicas — pelo fato de atingir as pessoas de forma rápida e ser muito mais em conta.

A segurança das pessoas, hoje, atrapalha os pesquisadores de campo quando estes vão até os condomínios e edifícios em muitos estados brasileiros. A difículdade existe desde a chegada do agente nos estabelecimentos, devido ao medo dos cidadãos em abrir a porta para estranhos até a negativa imposta por secretárias de prédios residenciais e comerciais. Pelo telefone o trabalho também não é fácil. Muitas vezes confundido com telemarketing, os pesquisadores escutam desaforos e, na maioria das vezes, o entrevistado não tem tempo disponível para responder as perguntas.

De acordo com profissionais do setor, a pesquisa on-line deve ser direcionada a produtos cujo público-alvo esteja na Internet. Desta forma, por meio da web, é possível direcionar o conteúdo para o público específico de um produto ou campanha.

A Internet veio para ficar e a cada dia conquista mais espaço entre as empresas e os consumidores. O fato da web não atingir a maioria da população brasileira, não diminui a sua importância para o mercado dinâmico de pesquisas. Sendo grande ou pequeno, o público não encontra desvantagens com pesquisas feitas pela Internet. Pelo contrário. Talvez para as empresas este problema exista, de fato.

O crescimento da pesquisa on-line no Brasil acontecerá à medida que a posse de computador e Internet nos lares aumentarem. Para descontinuar a forma de pesquisa convencional, o professor de Pesquisa de Mercado da ESPM, Victor Trujillo (FONTE: Thiago Terra, in www.mundodomarketing.com.br), afirma que só será possível quando a penetração da Internet for similar à da televisão. Sendo assim, de acordo com as previsões, teremos que responder ainda muitas pesquisas convencionais.

É inegável que haverá crescimento no uso da Internet entre os instrumentos de pesquisa de mercado. A esperança é que o público tenha acesso à rede e que a pesquisa saia de vez do papel e vá para o mundo virtual. O mercado de pesquisa on-line ganha terreno, os profissionais do setor apostam na Internet e os agentes de pesquisa torcem pela inclusão digital do seu trabalho.

#### 1.3 Métricas

As métricas avaliadas em uma pesquisa são classificações, comportamentais e atitudinais.

Eis aqui algumas métricas que podemos elencar para atingir nosso objetivo (GROSSNICKLE JOSHUA, 2001):

#### Classificação:

- Demográfica;
- Socioeconômica;
- Web gráfica;

#### Comportamental:

• Número de visitas a um determinado site;

- Número de compras online;
- Frequência de compras ou de participação em pesquisas online;

#### Atitudinal:

- Probabilidade de uso de determinado produto;
- Desejo específico;
- Escolha de marca;

Focando apenas em nosso objetivo, a abordagem em nossa pesquisa será apenas de "Classificação", procurando aperfeiçoar o tempo de resposta e consequentemente minimizar desistências.

#### 1.4 Escala Likert

Rensis Likert, nascido nos Estados Unidos em 1903, faleceu em 1981. Filho de engenheiro da Union Pacific Railroad seguiu inicialmente os passos do pai realizando seu treinamento profissional em engenharia. Influenciado pelos conflitos que pode observar durante a grande greve de 1922, resolveu estudas as organizações e o comportamento humano.

Rensis Likert recebeu seu B.A. em Sociologia pela Universidade de Michigan em 1926. Sua entrada precoce nestas áreas foi a base para muito do trabalho de Likert. O campo da sociologia na década de 1920 foi altamente experimental e incorporou muitos aspectos da psicologia moderna. Em 1932 recebeu seu Ph.D. em psicologia pela Universidade de Columbia. Por seu trabalho de tese, produzidos numa escala de pesquisa (Escalas de Likert), como um meio de medir atitudes, mostrando que é captada mais informação do que os métodos concorrentes. Em 1949, a instituição pioneira na pesquisa do comportamento humano nas organizações: o Institute for Social Research, da Universidade de Michigan. A sua obra teve grandes repercussões na teoria organizacional e no estudo da liderança. Likert baseou-se em inquéritos intensivos a empregados de companhias industriais, em que os interrogava sobre o comportamento dos seus supervisores. As respostas permitiram-lhe definir vários perfis ou estilos de liderança, que depois associou ao nível de desempenho das empresas.

Na década de 1960 Likert desenvolveu quatro sistemas de gestão, que descreveram a relação, o envolvimento e funções entre a direção e os subordinados em ambientes industriais. Os quatro sistemas são o resultado do estudo que ele fez com os supervisores altamente produtivos e os membros da equipe de uma companhia de seguros americana. Mais tarde, ele e Jane G. Likert revisaram os sistemas para aplicá-los nos contextos educativos.

Uma escala tipo Likert é composta por um conjunto de frases (itens) em relação a cada uma das quais se pede ao entrevistado para manifestar o grau de concordância desde o discordo totalmente (nível 0) até o concordo totalmente (nível 10). Avalia-se a atitude do sujeito somando ou calculando a média do nível selecionado para cada item.

Inicialmente Likert propôs um método de cálculo do resultado final por meio de uma média ponderada das respostas dadas, atribuindo em cada item um peso a cada nível de concordância expresso pelo entrevistado. Esses pesos são calculados assumindo uma

distribuição Normal e tomando como referência as frequências de resposta quando da aferição da escala (Ghiglione, 1993, p. 292).

No resumo do artigo de 1934, Likert, afirma: "É apresentado um método simples de atribuir scores em Escalas de Atitude de Thurstone, que não envolve o uso de um grupo de juízes e ainda se baseia em várias amostras para ser consistentemente mais fidedigno do que o método original." Não se muda de paradigma, mas apenas de forma de construção do instrumento de medida.

Forma de construção de uma escala de Likert:

Elaboração de uma lista de frases que manifestem opiniões radicais (claramente positivas ou negativas) em relação à atitude que se está a estudar, tendo o cuidado de cobrir as diferentes vertentes que se relacionam com o assunto.

Para validar a escala usa-se uma amostra representativa da população em que se pretende aplicar a escala de atitude. É pedido aos sujeitos que manifestem o seu grau de concordância com cada uma das frases numa escala de 3, 4, 5, 7 ou 11 pontos.

#### VANTAGENS:

Simples de construir Susceptível de ser criado com escala altamente confiável Fácil de ler e preencher

#### **DESVANTAGENS**:

Falta de reprodutibilidade

Ausência de unidimensionalidade

Difícil de demonstrar

Figura 2: Escala Likert - Fishbein and Ajzen, 1975

## Likert Scales

Please fill in the number that represents how you feel about the computer software you have been using

I am satisfied with it								
1	2	3	4	(5)				
Strongly Agree	Agree	Neither	Disagree	Strongly Disagree				
It is simple to	use							
1	2	3	4	(5)				
Strongly Agree	Agree	Neither	Disagree	Strongly Disagree				
It is fun to use	e							
1	2	3	4	(5)				
Strongly Agree	Agree	Neither	Disagree	Strongly Disagree				
It does every	thing I would e	expect it to do						
1	2	3	4	5				
Strongly Agree	Agree	Neither	Disagree	Strongly Disagree				
I don't notice	any inconsiste	encles as I use It	t					
1	2	3	4	(5)				
Strongly Agree	Agree	Neither	Disagree	Strongly Disagree				

Fonte: Fishbein and Ajzen, 1975.

Medidas em escala de Likert são fundamentalmente em nível ordinal de medida porque as respostas indicam um ranking.

Como o número de passos de escala é aumentada a partir de 2 até 20, o aumento da fiabilidade é muito rápida. Isso tende a estabilizar em cerca de 7, e depois de 11 passos, há pouco ganho em termos de fiabilidade em aumentar o número de passos (Nunally, 1978, citado por Neuman).

Curiosamente, Dyer (1995), "escalas de atitude não precisam ser factualmente precisas, elas simplesmente precisam refletir uma possível percepção da verdade".

De acordo com a declaração acima, a construção de uma escala de Likert de um conjunto de declarações, deve ser gerada relevante para a atitude (não necessariamente verdade). O número de escolhas na escala deve ser equilibrado para manter uma continuidade de declarações positivas e negativas com as quais o respondente é provável que concordar ou discordar, embora o número real de escolhas possa ser aumentado. Isso ajudará a evitar o

problema de viés e melhora a confiabilidade como quem responde 'concordo' o tempo todo vai aparecer para responder de forma inconsistente.

#### 2 AMOSTRA E COLETA

A coleta da amostra é outro ponto fundamental deste estudo. Por se tratar de uma pesquisa pela internet, temos pouca influência nas pessoas que responderão ou desistirão de responder cada questionário.

O trabalho de coleta online foi aleatório, respeitando atingir um número mínimo de respondentes, para que o experimento não ficasse prejudicado estatisticamente. Foi enviado a um mesmo grupo em dois momentos, os dois questionários objetos de estudo, com um período

A metodologia online foi adotada em função de sua dinâmica de entrega, seu baixíssimo custo operacional e também sua eficiência de consolidação dos dados.

Na sequência deste capitulo entenderemos melhor como foi realizada a coleta de dados neste estudo e como se apresentou a produtividade deste método.

#### 2.1 Amostra

O envio foi realizado aleatoriamente por meio de email, blogs e redes sociais, sem controle de perfil sócio demográfico, sexo ou qualquer outra classificação.

Um grande problema de produtividade nas pesquisas online é o spam, pois por meio de emails particulares, a entrega nem sempre é realizada. Além disso, existe muita desconfiança pela legitimidade e credibilidade da solicitação enviada.

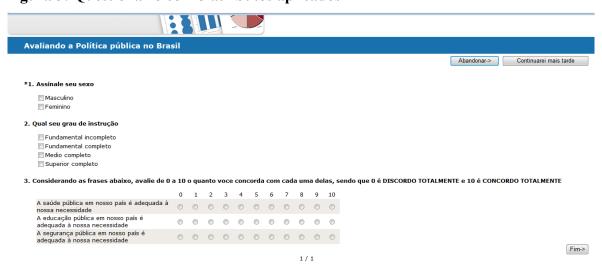
Neste estudo, no primeiro envio com os 3 atributos objeto de estudo, conseguimos que 156 abrissem o link do questionário, frente a 129 que efetivamente responderam. Uma taxa de produtividade de 83%.

No segundo envio, onde o questionário era substancialmente maior, tivemos 170 aberturas de link, contra 107 respostas efetivamente completadas, uma taxa de 64% de produtividade.

#### 2.2 Produtividade de Campo

Por meio das informações obtidas nesta coleta dos dados, temos nossa primeira inferência sobre o comportamento do entrevistado. O questionário de menor volume de perguntas teve um prazo de retorno menor e teve uma retenção do entrevistado maior

Figura 3: Questionário com 3 atributos aplicados



Fonte: Encuestafacil.com

Avaliando a Política pública no Brasil Abandonar-> Continuarei mais tarde \*1. Assinale seu sexo Masculino 2. Qual seu grau de instrução Fundamental incompleto Fundamental completo Medio completo Superior complete 3. Considerando as frases abaixo, avalie de 0 a 10 o quanto voce concorda com cada uma delas, sendo que 0 é DISCORDO TOTALMENTE e 10 é CONCORDO TOTALMENTE A mídia em nosso país visa a informação do cidadão 0 0 0 0 0 0 0 0 0 0 A corrupção em nosso país é uma das menores do o o o o o o o o mundo A política social em nosso país é adequada a 0 0 0 0 0 0 0 0 0 0 nossă necessidade
A presidente Dilma Roussef, está cumprindo o
prometido em campanha eleitoral
O transito de veículos nas vias públicas é seguro
As eleições em nosso país são justas
A saúde em nosso país é uma das melhores do
mundo. A segurança pública em nosso país é uma das melhores do mundo

A segurança pública em nosso país é uma das melhores do mundo

A violência tem diminuído a cada dia no Brasil

O investimento em esporte e cultura em nosso país é undo smaiores do mundo

A saúde pública em nosso país é adequada à nossa nacessidade

A qualidade dos serviços públicos tem melhorado a cada ano no Brasil

A educação pública em nosso país é adequada à a Cada ano no Brasii A educação pública em nosso país é adequada à O desemprego tem diminuído a cada dia no Brasil incia social em nosso país é adequada a 0 0 0 0 0 0 0 0 0 0 nossa necessidade Fim->

Figura 4: Questionário com 17 atributos aplicados

Fonte: Encuestafacil.com

Em pesquisas com grandes escalas de atributos de concordância, é comumente utilizado o rodízio de perguntas, para que a posição do atributo seja uma variável isolada do modelo. Neste estudo, foi adotada a abordagem sem rodízio, pois a ferramenta não permitia essa funcionalidade.

Importante enfatizar, que os canais utilizados para envio dos questionários foram os mesmos, onde foi enviado a um grupo restrito de pessoas, de aproximadamente 700 pessoas, e, portanto tivemos um retorno de aproximadamente 20%, entre pessoas que responderam apenas um deles ou os dois. Buscando a discrição e maior conforto dos participantes, não foi exigida identificação do entrevistado.

#### 3 ANÁLISE DESCRITIVA

Após a coleta dos dados, analisamos descritivamente os resultados abordados nos questionários estudados. Esta etapa é de suma importância na análise primária dos resultados, podendo observar possíveis tendências e características particulares de perfil e paramétrica.

Trabalharemos na análise em duas frentes, uma para o 1º Questionário, com 3 atributos e uma para o 2º Questionário, com 17 atributos.

Na sequência, temos a distribuição por sexo e escolaridade e as médias atingidas com os atributos avaliados, além da análise de comparabilidade descritiva entre os atributos.

Por fim, neste capítulo, entenderemos teoricamente, quais os próximos passos analíticos foi adotado no estudo, como o teste das médias e qual método utilizado.

#### 3.1 Questionário 1:

#### Três Atributos:

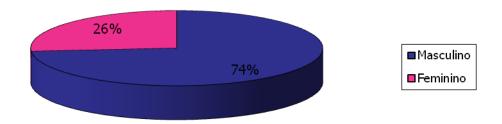
- "A saúde pública em nosso país é adequada à nossa necessidade";
- "A segurança pública em nosso país é adequada à nossa necessidade";
- "A educação pública em nosso país é adequada à nossa necessidade"

#### <u>Amostra</u>

Foram entrevistadas, no sistema de autopreenchimento, 129 pessoas.

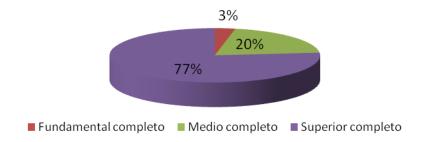
São formadas por usuários de internet, redes sociais com conhecimento suficiente para o autopreenchimento. Na sua maioria, está formado por homens. Como o sexo não foi definido como variável discriminante no resultado, este desequilíbrio não influenciará no teste.

Figura 5: Gráfico da amostra por sexo – Fase 1



Em relação ao grau de instrução, 77% dos entrevistados possui nível superior, sendo que outros 20% possui ao menos o ensino médio.

Figura 6: Gráfico da amostra por grau de escolaridade – Fase 1



## Avaliação dos Atributos:

Os três Atributos avaliados neste questionário 1, tiveram um comportamento similar em relação a média e variabilidade, entre eles.

Figura 7: Analise descritiva do Questionário 3 atributos

Atributos (escala 0 a 10)	Média	Desvio Padrão	Mediana	Moda
1. "A saúde pública em nosso				
país é adequada à nossa	2,13	1,8	2	0
necessidade";				
2. "A segurança pública em				
nosso país é adequada à nossa	2,18	1,9	2	0
necessidade";				
3. "A educação pública em				
nosso país é adequada à nossa	2,01	1,8	2	0
necessidade				

#### 3.2 Questionário 2

#### Os 17 Atributos:

- "As leis tributárias no Brasil são justas"
- "A mídia em nosso país visa a informação do cidadão"
- "A corrupção em nosso país é uma das menores do mundo"
- "A política social em nosso país é adequada a nossa necessidade"
- "O transporte público no Brasil é adequado à nossa necessidade"
- "A segurança pública em nosso país é adequada à nossa necessidade"
- "A presidente Dilma Roussef, está cumprindo o prometido em campanha eleitoral"
- "O transito de veículos nas vias públicas é seguro"
- "As eleições em nosso país são justas"
- "A saúde em nosso país é uma das melhores do mundo"
- "A segurança pública em nosso país é uma das melhores do mundo"
- "A violência tem diminuído a cada dia no Brasil"
- "O investimento em esporte e cultura em nosso país é um dos maiores do mundo"
- "A saúde pública em nosso país é adequada à nossa necessidade"
- "A qualidade dos serviços públicos tem melhorado a cada ano no Brasil"
- "A educação pública em nosso país é adequada à nossa necessidade"
- "O desemprego tem diminuído a cada dia no Brasil"
- "A previdência social em nosso país é adequada a nossa necessidade"

Foram entrevistadas, no sistema de autopreenchimento, 107 pessoas. São formadas por usuários de internet, redes sociais com conhecimento suficiente para a participação. Na sua maioria, está formado por homens. Como o sexo não foi definido como variável discriminante no resultado, este desequilíbrio não influenciará no teste.

Figura 8: Gráfico da amostra por sexo – Fase 2

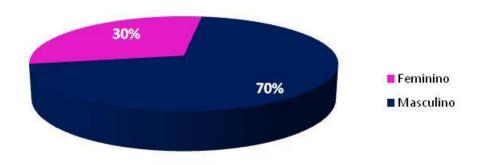


Figura 9: Gráfico da amostra por grau de escolaridade – Fase 2

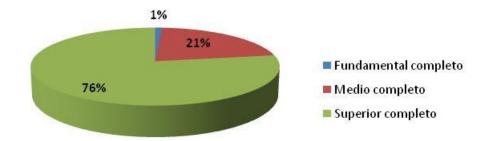


Figura 10: Analise descritiva do Questionário 17 atributos

Atributos (Escala 0 a 10)	Média	Desvio Padrão	Mediana	Moda
A previdência social em nosso país é adequada a nossa necessidade	1.88	1.79	2	0
A política social em nosso país é adequada a nossa necessidade	2.20	2.01	2	0
As eleições em nosso país são justas	5.07	3.22	5	5
A saúde em nosso país é uma das melhores do mundo	1.64	1.94	1	0
O investimento em esporte e cultura em nosso país é um dos maiores do mundo	1.78	1.81	1	0
A segurança pública em nosso país é uma das melhores do mundo	1.35	1.71	1	0
A saúde pública em nosso país é adequada à nossa necessidade	1.37	1.46	1	0
A corrupção em nosso país é uma das menores do mundo	0.83	1.69	0	0
O transporte público no Brasil é adequado à nossa necessidade	1.58	1.63	1	0
As leis tributárias no Brasil são justas	1.13	1.57	0	0
A segurança pública em nosso país é adequada à nossa necessidade	1.61	1.50	2	0
O transito de veículos nas vias públicas é seguro	2.47	2.30	2	0
A presidente Dilma Roussef, está cumprindo o prometido em campanha eleitoral	3.77	2.31	4	5
O desemprego tem diminuído a cada dia no Brasil	5.07	2.46	5	5
A violência tem diminuído a cada dia no Brasil	1.97	2.05	1	0
A qualidade dos serviços públicos tem melhorado a cada ano no Brasil	2.49	2.05	2	0
A mídia em nosso país visa a informação do cidadão	2.61	2.09	2	2
A educação pública em nosso país é adequada à nossa necessidade	1.50	1.70	1	0

#### 3.3 Comparabilidade

De forma descritiva, os primeiros resultados do teste em estudo estão mostrando que na fase 1, as médias são ligeiramente maiores que na fase 2, considerando os três atributos pesquisados. Além disso, a variabilidade na Fase 2, é menor que na fase 1, corroborando com a hipótese de que em um questionário com mais avaliações, o entrevistado tende a ser mais constante em suas respostas.

Segundo Gunther, H. (1999), um questionário bem estruturado contribui para a redução do desgaste físico e mental do respondente, garantindo que os temas de interesse do pesquisador sejam tratados de forma cadenciada numa ordem fluida capaz de manter o entrevistado motivado a responder. Para tanto, é preciso foco no objetivo da pesquisa, nas perguntas que o pesquisador espera responder por meio dela, tendo a clara noção do que está incluído em cada ítem do instrumento.

Fowler (1998, p.344) define um bom item como aquele que gera respostas fidedignas e válidas. Apresenta cinco características básicas:

a. A pergunta precisa ser compreendida consistentemente

- b. A pergunta precisa ser comunicada consistentemente
- c. As expectativas quanto à resposta dada precisam ser clara para o respondente
- d. A menos que estejam verificando conhecimento, os respondentes devem ter toda informação necessária.
- e. Os respondentes precisam estar dispostos a responder

O público pesquisado nas duas fases é formado em sua maioria pelo sexo masculino com superior completo.

Na fase 1, temos um comportamento aparentemente distinto entre os sexos.

Figura 11: Analise descritiva por sexo – Fase 1

Atributes Food 1	Assinale seu sexo			
Atributos - Fase 1	Feminino	Masculino		
A saúde pública em nosso país é adequada à nossa necessidade	1.76	2.26		
A segurança pública em nosso país é adequada à nossa necessidade	1.85	2.29		
A educação pública em nosso país é adequada à nossa necessidade	1.74	2.11		

Na fase 2, o mesmo evento ocorre, exceto no atributo referente a educação pública, onde os resultados são praticamente iguais.

Figura 12: Analise descritiva por sexo – Fase 2

Atributes Fore 2	Assinale seu sexo			
Atributos - Fase 2	Feminino	Masculino		
A saúde pública em nosso país é adequada à nossa necessidade	.97	1.55		
A segurança pública em nosso país é adequada à nossa necessidade	1.19	1.79		
A educação pública em nosso país é adequada à nossa necessidade	1.56	1.48		

#### 3.4 Teste de Médias

Para definirmos a significância estatística das médias estudadas nas duas fases, vamos realizar o test t de Student para médias.

O teste de Student, ou simplesmente *teste t* o método mais utilizado para se avaliarem as diferenças entre as médias de dois grupos. Para estudar o efeito de um experimento, muitas vezes comparam-se pares de indivíduos.

A estatística t foi introduzida em 1908 por William Sealy Gosset, químico da cervejaria Guinness em Dublin, Irlanda ("student" era seu pseudônimo). Gosset havia sido contratado devido à política inovadora de Claude Guinness de recrutar os melhores graduados

de Oxford e Cambridge para os cargos de bioquímico e estatístico da indústria Guinness. Gosset desenvolveu o Teste t como um modo barato de monitorar a qualidade da cerveja tipo stout. Ele publicou o Teste t na revista acadêmica Biometrika em 1908, mas foi forçado a usar seu pseudônimo pelo seu empregador, que acreditava que o fato de usar estatística era um segredo industrial. De fato, a identidade de Gosset não foi reconhecida por seus colegas estatísticos.

Se forem feitas inúmeras amostras de tamanho n a partir da mesma população e se fossem tiradas as médias de uma variável dessa população que possui uma distribuição normal, a distribuição dessas inúmeras médias seguiria uma distribuição t de Student. Por exemplo, imaginemos que a altura das pessoas segue uma distribuição normal. Se selecionarmos diversas amostras aleatórias de 100 pessoas e calculássemos a média da altura das pessoas de cada amostra, essa média da altura das pessoas seguirá uma distribuição t de Student.

Perceba que, na distribuição t de Student, valores muito baixos ou muito altos tem menor probabilidade de ocorrer, indicando que é menos provável que a média de uma amostra apresente valores muito distantes da média da população.

O formato da distribuição t de Student depende do número de graus de liberdade. Quanto maior o número de graus de liberdade, mais "concentrada" é a distribuição. Para valores muito grandes de graus de liberdade, a distribuição t de Student se aproxima da distribuição normal.

O Teste t consiste em formular uma hipótese nula e consequentemente uma hipótese alternativa, calcular o valor de t conforme a fórmula apropriada (abaixo) e aplicá-lo à função densidade de probabilidade da distribuição t de Student medindo o tamanho da área abaixo dessa função para valores maiores ou iguais a t. Essa área representa a probabilidade da média dessa(s) amostra(s) em questão ter(em) apresentado o(s) valor(es) observado(s) ou algo mais extremo. Se a probabilidade desse resultado ter ocorrido for muito pequena, podemos concluir que o resultado observado é estatisticamente relevante. Essa probabilidade também é chamada de p-valor ou valor p. Consequentemente, o nível de confiança  $\alpha$ é igual a 1 - p-valor.

Normalmente é usado um "ponto de corte" para o p-valor ou para o nível de confiança para definir se a hipótese nula deve ser rejeitada ou não. Se o p-valor for menor que esse "ponto de corte", a hipótese nula é rejeitada. Caso contrário, a hipótese nula não é rejeitada.

É comum que sejam usados os "pontos de corte" para p-valor 0,1%, 0,5%, 1%, 2% ou 5%, fazendo com que os níveis de confiança sejam, respectivamente, 99,9%, 99,5%, 99%,

98% ou 95%. Caso seja usado o p-valor 5% como "ponto de corte" e a área abaixo da função densidade de probabilidade da distribuição t de Student seja menor do que 5%, pode-se afirmar que a hipótese nula é rejeitada com nível de confiança de 95%.

Note que não rejeitar a hipótese nula não é a mesma coisa que afirmar que a hipótese alternativa é válida com o mesmo nível de confiança. Isso seria uma interpretação incorreta do teste.

Dependendo da definição da hipótese nula, deve ser usado uma ou duas caudas da distribuição t de Student na avaliação do teste. Por exemplo, se a hipótese nula for  $\bar{x} \leq \mu_{0e}$  a hipótese alternativa  $\bar{x} > \mu_{0}$ , o teste deve ser feito somente para valores maiores do que  $t_{e}$ , portanto, ao consultar a função densidade de probabilidade da distribuição t de Student, devese considerar somente a área superior a t, ou seja, somente uma das "caudas" da distribuição.

Por outro lado, se a hipótese nula for  $\bar{x}=\mu_0$  e, consequentemente, a hipótese alternativa  $\bar{x}\neq\mu_0$ , teríamos que avaliar ao mesmo tempo a possibilidade de  $\bar{x}<\mu_0$  e de  $\bar{x}>\mu_0$ . Para isso, ao consultar a função densidade de probabilidade da distribuição t de Student , devem ser consideradas as áreas abaixo da curva para valores superiores a t e inferiores a -t, ou seja, as duas "caudas" da distribuição. Como a distribuição é simétrica, os tamanhos dessas áreas são iguais.

#### 3.5 Aplicação do teste t para amostras independentes

Estes testes se aplicam a planos amostrais onde se deseja comparar dois grupos independentes. Esses grupos podem ter sido formados de duas maneiras diferentes:

- a) Extraiu-se uma amostra da população A e outra amostra da população B;
- b) Indivíduos da mesma população foram alocados aleatoriamente a um dos dois tratamentos em estudo.

Diferente do caso de dados pareados, não se exige que as amostras tenham o mesmo tamanho.

Quando as suposições paramétricas são atendidas, o teste t de Student para amostras independentes é mais adequado para comparação das médias dos grupos.

#### **Exemplos**

'Student' Test t é uma das técnicas mais utilizadas para testar uma hipótese sobre a base de uma diferença entre as médias das amostras. Explicada em termos genéricos, o teste t

determina uma probabilidade de que duas populações são o mesmo no que diz respeito à variável testada.

Por exemplo, suponha que você coletou dados sobre as alturas de basquete masculino e jogadores de futebol, e compararam às amostras utilizando o teste t. A probabilidade de 0,4 significaria que há uma significância de 40% que você não consegue distinguir um grupo de jogadores de basquete de um grupo de jogadores de futebol por altura. Se você calcular uma probabilidade de 0,05 ou menos, então você pode rejeitar a hipótese nula (ou seja, você pode concluir que os dois grupos de atletas podem ser distinguidos pela altura).

Para o teste t de Student para amostras independentes que não têm que ter o mesmo número de pontos de dados em cada grupo, temos de assumir que a população segue uma distribuição normal.

O teste t pode ser realizado sabendo apenas as médias, desvio padrão e número de pontos de dados. A amostra de teste t duas produz uma estatística t, no qual X-bar, é claro, é a média da amostra, e S é o desvio padrão da amostra. O denominador é uma medição do erro experimental nos dois grupos combinados. Quanto maior a diferença entre as médias, mais confiante você está nos dados. Assim, quanto mais alto o valor de T, maior a confiança de que há uma diferença.

Segundo Richard Lowry, 1999-2012, Este é provavelmente o teste estatístico mais utilizado de todos os tempos, e certamente o mais conhecido. É simples, direto, fácil de usar e adaptável a uma ampla gama de situações. Na caixa de ferramentas estatística deve sempre estar com ele. A sua utilidade é ocasionada pelo fato de que a pesquisa científica, muitas vezes analisa os fenômenos da natureza duas variáveis de cada vez, com um olho para responder à pergunta básica: Será que estas duas variáveis estão relacionadas? Se alterar o nível de um, vamos assim alterar o nível do outro? Ou, em alternativa: Se examinarmos dois níveis diferentes de uma variável vamos associar com níveis diferentes de outro?

#### 4 RESULTADOS

Neste capítulo, foram realizados testes estatísticos de média, para avaliarmos se existe realmente diferença significativa entre os questionários e se esta diferença se aplica aos 3 atributos ou apenas um deles.

A seguir teremos o resultado para cada um dos 3 atributos avaliados. Saberemos se todos tem o mesmo comportamento ou se são distintos entre si.

#### 4.1 Atributo 1 "A saúde pública em nosso país é adequada à nossa necessidade"

Vamos avaliar se existe diferença significativa entre as médias estudadas, do atributo em questão, nas duas avaliações realizadas. No GRUPO 1 considerando apenas os 3 atributos objetos de estudo e no GRUPO 2, considerando os 17 atributos.

Abaixo, uma análise gráfica do comportamento das duas avaliações, neste atributo.

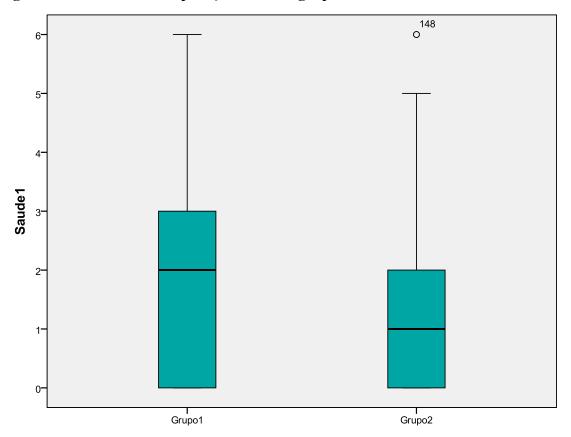


Figura 13: Box Plot – Comparação entre os grupos – Atributo Saúde

Fonte: Ferramenta estatística SPSS

Pelo gráfico acima, notamos uma diferença aparente entre as medidas centrais e na dispersão entre os dois grupos estudados.

#### **Teste t Student – Amostras Independentes**

O teste a seguir, visa uma conclusão, baseado nas seguintes hipóteses:

H0: As medias entre os atributos nas duas aplicações são iguais

H1: As medias entre os atributos nas duas aplicações não são iguais

Utilizamos o software SPSS, como ferramenta para os cálculos desta análise.

**Group Statistics** 

	Grupo Saúde	N	Mean	Std. Deviation	Std. Error Mean
Saúde	1	129	2.13	1.765	.155
	2	107	1.37	1.457	.141

Aqui, vemos novamente, que existe uma diferença considerável entre as médias, aplicadas nas diferentes formas. A média do Grupo 1 é aproximadamente 35% maior que no grupo 2. Mesmo sob estas evidências, ainda não podemos concluir que existe diferença significativa entre os atributos neste experimento.

Teste de Amostras Independentes

Independent Samples Test

	_	Levene's Test Varia	for Equality of nces				t-test for Equality	of Means	•	•
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Saude1	Equal variances assumed	4.181	.042	3.550	234	.000	.758	.213	.337	1.179
	Equal variances not assumed			3.614	233.996	.000	.758	.210	.345	1.171

Neste estudo, vamos assumir um nível de significância padrão de 95% de confiabilidade.

Pelo teste acima, de igualdade de médias e variâncias, conseguimos identificar por meio do P-valor (Sig./ Sig. (2-tailed)), ao nível de sig. Assumido, que tanto o teste de igualdade de variância de Levene's e o Teste t de igualdade de médias, que devemos rejeitar a hipótese 1, de igualdade das médias.

Rejeitamos H0, quando o P-valor for menor ou igual a 0,05, ao nível de significância assumido, caso contrário, poderíamos dizer que Não rejeitamos a Hipótese H0, assumindo igualdade entre as médias estudadas.

Para a primeira variável em estudo, podemos concluir que existe diferença significativa na média, entre os dois formatos de aplicação. Vale ressaltar, o posicionamento do atributo no questionário, onde "A Saúde pública em nosso país é adequado à nossa necessidade" é aplicado na primeira posição do questionário 1, enquanto no questionário 2, ele é aplicado na posição 7.

# 4.2 Atributo 2 "A segurança pública em nosso país é adequada à nossa necessidade"

Agora avaliaremos se existe diferença significativa entre as médias estudadas, do atributo em questão, nas duas avaliações realizadas. No GRUPO 1 considerando apenas os 3 atributos objetos de estudo e no GRUPO 2, considerando os 17 atributos.

Análise gráfica do comportamento das duas avaliações, neste atributo.

Sedinanca

Grupo1

Grupo2

Figura 14: Box Plot – Comparação entre os grupos – Atributo Segurança

Fonte: Ferramenta estatística SPSS

Pelo gráfico de Box Plot, podemos notar uma leve similaridade na mediana, porém com uma dispersão maior no grupo 1 entre as notas mencionadas no estudo.

#### **Teste t Student – Amostras Independentes**

Novamente, baseado nas seguintes hipóteses:

H0: As medias entre os atributos nas duas aplicações são iguais

H1: As medias entre os atributos nas duas aplicações não são iguais

Utilizamos o software SPSS, como ferramenta para os cálculos desta análise.

**Group Statistics** 

	Grupo Segurança	N	Mean	Std. Deviation	Std. Error Mean
Segurança	1	129	2.18	1.873	.165
	2	107	1.61	1.497	.145

Aqui, vemos que existe uma diferença considerável entre as médias, aplicadas nas diferentes formas. A média do Grupo 1 é aproximadamente 35% maior que no grupo 2. Pelos resultados atingidos no atributo passado, temos uma tendência à mesma conclusão.

Teste de Amostras Independentes

Independent Samples Test

		Levene's Test Varia	t-test for Equality of Means							
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Seguranca1	Equal variances assumed	3.129	.078	2.549	234	.011	.571	.224	.130	1.012
	Equal variances not assumed			2.602	233.699	.010	.571	.219	.139	1.003

Assumindo um nível de significância padrão de 95% de confiabilidade, podemos notar que não existe uma diferença significativa pelo Teste de variâncias de Levene's. Porém, se considerarmos um nível de confiança um pouco menos preciso, de 90%, já conseguiríamos assumir que existe diferença entre as variâncias nos dois formatos de avaliação.

Para ratificar a diferença entre as médias coletadas, vamos analisar agora o Teste t de igualdade de médias de amostras independentes. Podemos novamente notar pelo p-valor

(0,011), que ao nível de significância de 95%, existe diferença significativa entre as médias observadas.

Portanto, rejeitamos H0, para a variável em estudo. Podemos concluir que existe diferença significativa na média, entre os dois formatos de aplicação. Outro indicador que nos confirma essa diferença, é o Intervalo de Confiança de igualdade das médias. O valor inferior é maior que 0 (zero), o que nos ratifica a conclusão de que existe diferença significativa entre as médias. Isto acontece quando o valor inferior é maior que 0 ou o superior é menor que 0.

Ressaltando, novamente, o posicionamento do atributo no questionário, onde "A Segurança pública em nosso país é adequada à nossa necessidade" é aplicado na segunda posição do questionário 1, enquanto no questionário 2, ele é aplicado na posição 11.

# 4.3 Atributo 3 "A educação pública em nosso país é adequada à nossa necessidade"

Por fim, avaliaremos se existe diferença significativa entre as médias estudadas, do atributo sobre a educação pública, nas duas avaliações realizadas. No GRUPO 1 considerando apenas os 3 atributos objetos de estudo e no GRUPO 2, considerando os 17 atributos.

Análise gráfica do comportamento das duas avaliações, neste atributo.

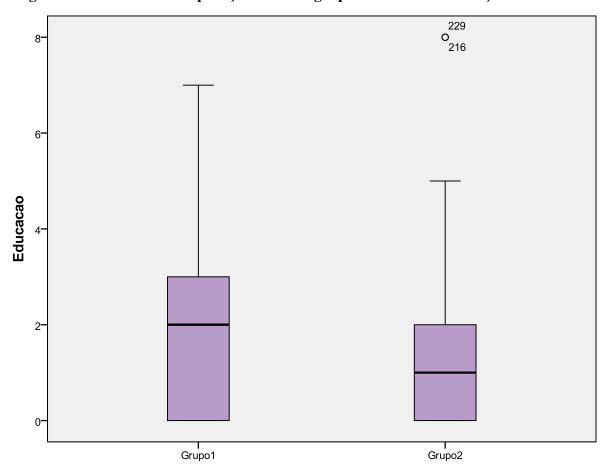


Figura 15: Box Plot – Comparação entre os grupos – Atributo Educação.

**Fonte:** Ferramenta estatística SPSS

Neste atributo, podemos notar que as dipersões entre os dois grupos são menores, e similares entre si.

#### **Teste t Student – Amostras Independentes**

Considerando as hipóteses para igualdade de médias:

H0: As medias entre os atributos nas duas aplicações são iguais

H1: As medias entre os atributos nas duas aplicações não são iguais

Utilizamos o software SPSS, como ferramenta para os cálculos desta análise.

**Group Statistics** 

	Grupo Educação	N	Mean	Std. Deviation	Std. Error Mean
Educação	1	129	2.01	1.783	.157
	2	107	1.50	1.701	.164

Aqui, vemos que existe uma diferença entre as médias aplicadas nas diferentes formas. A média do Grupo 1 é aproximadamente 30% maior que no grupo 2. Pelos resultados atingidos no atributo passado, temos uma tendência à mesma conclusão.

Teste de Amostras Independentes

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidenc Differ	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Educacao1	Equal variances assumed	.778	.379	2.203	234	.029	.503	.228	.053	.953
	Equal variances not assumed			2.213	229.438	.028	.503	.227	.055	.951

Assumindo um nível de significância padrão de 95% de confiabilidade, podemos notar que não existe uma diferença significativa pelo Teste de variâncias de Levene's. Notando o P-valor desta análise, com o valor de 0.379, podemos considerar com maior precisão que não existe diferença significativa na variabilidade encontrada nos dois grupos.

Já no teste t de Student, de igualdade de **médias**, novamente, assim como nos dois primeiros atributos, notamos pelo p-valor encontrado (0.029) que ao nível de significância de 95% existe diferença significativa entre as médias estudadas.

Além disso, pelo intervalo de confiança encontrado, notamos que ele não compreende o valor 0, confirmando nossa Hipótese 1, de que não existe igualdade entre as médias estudadas

Portanto, rejeitamos H0, para a variável em estudo. Podemos concluir que existe diferença significativa na média, entre os dois formatos de aplicação.

O posicionamento do atributo no questionário, onde "A Segurança pública em nosso país é adequada à nossa necessidade" é aplicado na terceira posição do questionário 1, enquanto no questionário 2, ele é aplicado na posição 17.

## **CONCLUSÃO**

Como ponto de partida, nosso principal objeto de estudo foi avaliar o impacto na avaliação de atributos de concordância em duas formas de distinta de avaliação.

Na criação de um questionário, mais precisamente considerando as notas de concordância de uma escala likert, precisamos ter muito cuidado com o que queremos avaliar. Hoje em dia, com a necessidade de informação do marketing, acabamos por sobrecarregar um questionário tornando inviável manter a qualidade da resposta obtida. Vimos neste estudo, que existe sim uma diferença entre diferentes métodos e estes métodos, podem influenciar drasticamente no resultado de uma mesma resposta.

Olhando de forma mais cautelosa para nossos resultados, conseguimos observar que no primeiro experimento (questionário 1), as notas médias dos 3 atributos avaliados apresentam notas mais altas. No segundo experimento (questionário 2), o entrevistado acabou avaliando de forma mais crítica os atributos relacionados à educação, saúde e segurança pública, influenciado pelos outros temas que avaliamos no mesmo questionário, nas mesma bateria. Ao responder sobre performance da atual presidente, Midia e qualidade de nosso processo eleitoral, as notas tendem a ser baixas, mas consideravelmente mais altas que as 3 de nosso objeto de estudo, pois são teoricamente menos problemáticas na nossa sociedade.

Conseguimos concluir alguns fatores que influenciaram diretamente no resultado deste experimento. A primeira delas, é em relação ao **tempo**. No questionário 2, o tempo médio de autopreenchimento do questionário foi de 4:42 minutos, enquanto no questionário 1, apenas 2:31 minutos. Esta diferença entre ambos pode ocasionar uma fadiga no entrevistado e ele tende a pensar pouco nas respostas e ler os atributos com menor atenção. Este ponto ficou mais nítido na produtividade dos dois questionários, onde o primeiro está na casa de 16% de desistência após abertura do link versus 36% de desistência no segundo questionário (% de entrevistados que fecharam o questionário antes de iniciar a resposta).

O segundo ponto observado como influente, na diferença entre as médias, e este podeé até mais importante que o primeiro, é o fator **influência** entre os atributos. Ao passar item a item do questionário, ele indubitavelmente vai comparando uma avaliação com as anteriores e a atual acaba sofrendo influência na avaliação em relação a anterior. Se ele dá uma nota baixa para algum tema que julga crítico, e na pergunta seguinte ele nota que o tema seguinte é mais crítico que o anterior, o entrevistado acaba dando uma nota menor que a avaliação anterior, mesmo as duas sendo relativamente baixas.

Considerando que na bateria 2 as notas foram mais baixas que na bateria 1, a tese de influência ganha mais força.

Esta hipótese nos faz pensar melhor em como estruturar nossas escalas Likert em um questionário quantitativo. Devemos priorizar a avaliação de vários itens ao mesmo tempo, considerando como ganho a quantidade de informações ou queremos priorizar o que realmente vai impactar em nosso plano de ação?

Outro ponto importante a destacar é a posição no questionário de cada um dos atributos. Podemos inferir que quanto mais para o final ele for abordado, maior é a influência que os outros atributos incidem sobre ele. Podemos perceber isso especialmente neste estudo, pois não adotamos o esquema de rodízio.

Sabemos que no mercado de pesquisa, dificilmente conseguimos trabalhar como no questionário 1. Normalmente, os questionários são longos e carregados e na maioria das vezes acabamos tendo uma colinearidade entre as avaliações, pois em meio há tantas avaliações, o entrevistado acaba tendo uma escassez mental de julgamentos, tendo por uniformizar suas avaliações e no final do estudo não conseguimos destacar os pontos que precisaremos nos aprofundar.

Neste estudo, em relação aos temas escolhidos, foi priorizado temas de conhecimento público e principalmente objeto de muita crítica social. Talvez, se perguntássemos apenas sobre o "Serviço público" oferecido à população, independente de qual serviço conseguiria reduzir ainda mais nossa bateria.

Finalmente, devemos repensar a cada plano de marketing como e onde devemos aplicar nossos objetivos de pesquisa em um questionário. Vimos que perguntas idênticas apresentam resultados diferentes o que torna nosso estudo mais vulnerável, dependendo de nosso foco.

## REFERÊNCIAS

CHURCHILL, G.A.; IACOBUCCI, D. Marketing Research Methodological Foundations. 9<sup>th</sup> ed. London: South Western, 2010.

FOWLER, F. J. **Design and evaluation of survey questions:** Handbook of applied social research methods. Thousand Oaks: Sage, 1988. (pp. 343-374).

GROSSNICKLE, J.; RASKIN, O. **The Handbook of Online Marketing Research**. 1<sup>st</sup> ed. New York: McGraw-Hill Companies, 2001.

HAIR, J. et al. **Fundamentos de Métodos de Pesquisa em Administração**. 5<sup>a</sup> ed. Porto Alegre: Bookman, 2005. pp 158-159.

LIKERT, R. Novos Padrões de Administração. 2ª ed. São Paulo: McGraw-Hill, 1971.

MALHOTRA, N. K. **Pesquisa de Marketing:** Uma orientação aplicada. 4ª ed. Porto Alegre: Bookman, 2006 pp 266-276.

MALHOTRA, N. K. **Pesquisa de Marketing:** Uma orientação aplicada. 4ª ed. Porto Alegre: Bookman, 2006. pp 290-291.

MARTINEZ, L. F. A. **Análise de dados com SPSS primeiros passos**. 3ª ed. Lisboa: Escolar Editora, 2007.

MORETTIN, P.A.; BUSSAB, W. O. Estatística Básica. 5ª ed. Rio de Janeiro: Saraiva, 2004.

NOETHER, G. E. **Introdução a Estatística**. 2ª ed. Trad Flavio Rodrigues e Lisbeth K. Cordani. Rio de Janeiro: Guanabara Dois, 1983.

PADUA, E.M.M. **Metodologia de Pesquisa Abordagem teórico Pratica**, 13<sup>a</sup> ed. Campinas: Papirus, 2007.

PsycINFO. [ Journal Article ]. A technique for the measurement of attitudes. LIKERT, R. Archives of Psychology No. 140.1932

REA, L.M.; PARKER R.A. **Metodologia de Pesquisa do Planejamento a Execução.** Ed. São Paulo: Pioneira, 2000.

SPIEGEL, M. R. Estatística. 3ª ed. Trad Pedro Consentino. São Paulo: Mc Graw-Hill do Brasil, 1979.