

**UNIVERSIDADE FEDERAL DE SÃO CARLOS**  
**CENTRO DE EDUCACAO E CIENCIAS HUMANAS (CECH)**  
**CURSO DE PSICOLOGIA**

**Desejabilidade social e avaliação da empatia em testes de autorrelato: Estudo de uma  
nova escala**

**IURI LIVRAMENTO TONIATO**

São Carlos - SP

2025

IURI LIVRAMENTO TONIATO

**Desejabilidade social e avaliação da empatia em testes de autorrelato: Estudo de uma  
nova escala**

Trabalho de conclusão de curso apresentado  
ao curso de psicologia da Universidade  
Federal Carlos, para obtenção do título de  
bacharel em Psicologia

Orientador: Fabiano Koich Miguel

São Carlos - SP

2025

## **Agradecimentos**

Aos meus pais, Josefina e Dail, que sempre apoiaram e acreditaram nos meus sonhos. Agradeço por darem asas aos meus voos e por serem a terra firme para meus pousos. Vocês sempre serão o meu lar. A todos os professores que participaram da minha formação educacional, cada palavra, cada conselho, crítica e elogio me guiaram a trilhar o caminho que sigo hoje. Agradeço ao meu namorado, Lucas R., por seu apoio e companheirismo. Ao meu orientador, Fabiano Miguel, por sua paciência e didática na construção desta pesquisa. Às minhas amigas, Paula R., Luana B. e Luana M., pelos inúmeros cafés com monografia e pelas orientações compartilhadas sobre escrita e vida acadêmica. Agradeço a mim por meu empenho e dedicação em conquistar os espaços que sonhei para mim.

## Resumo

A avaliação da empatia, uma importante habilidade social, frequentemente se baseia em testes de autorrelato, que são vulneráveis à desejabilidade social (DS). Esse viés, caracterizado pela tendência dos indivíduos a responder de forma socialmente aceitável, pode distorcer as respostas e comprometer a precisão dos resultados, especialmente em medidas de empatia. O presente estudo examinou o impacto da DS na avaliação da empatia por meio do Questionário Online de Empatia (QoE) e testou um novo formato de escala (QoE-2B), focado na frequência de comportamentos empáticos, visando minimizar esse viés. Participaram do estudo 1005 indivíduos, que responderam versões da escala Likert tradicional (QoE-2A) ou a nova escala proposta (QoE-2B), além da Escala de Desejabilidade Social de Marlowe-Crowne (EDSMC-20) e do Big Five Inventory 2 (BFI-2-S). Os resultados indicaram que a QoE-2A apresentou uma correlação positiva e significativa com a DS ( $r=0.35, p<0.01$ ), enquanto a QoE-2B não demonstrou correlação significativa com a DS ( $r=-0.01$ ). Em relação aos traços de personalidade, a QoE-2A correlacionou-se conforme esperado com o BFI (e.g., Agradabilidade,  $r=0.67$ ), ao passo que a QoE-2B não mostrou correlações significativas com nenhuma das dimensões do BFI. Conclui-se que o novo formato da escala (QoE-2B) é promissor na mitigação do viés da desejabilidade social, oferecendo uma avaliação potencialmente mais fidedigna da empatia e levantando questões sobre os diferentes aspectos que cada formato pode medir.

Palavras-chave: desejabilidade social, empatia, testes de autorrelato, avaliação psicológica, personalidade.

## Abstract

The assessment of empathy, a crucial social skill, frequently relies on self-report measures, which are vulnerable to social desirability bias (SDB). This bias, characterized by individuals' tendency to respond in a socially acceptable manner, can distort responses and compromise the accuracy of results, particularly in empathy measures. The present study examined the impact of SDB on empathy assessment using the Questionário Online de Empatia (QoE) and tested a new scale format (QoE-2B), focused on the frequency of empathic behaviors, aiming to minimize this bias. A total of 1005 individuals participated in the study, responding to either the traditional Likert scale version (QoE-2A) or the newly proposed scale (QoE-2B), in addition to the Marlowe-Crowne Social Desirability Scale (EDSMC-20) and the Big Five Inventory 2 (BFI-2-S). The results indicated that QoE-2A showed a positive and significant correlation with SDB ( $r=0.35, p<0.01$ ), whereas QoE-2B did not demonstrate a significant correlation with SDB ( $r=-0.01$ ). Regarding personality traits, QoE-2A correlated as expected with the BFI (e.g., Agreeableness,  $r=0.67$ ), while QoE-2B did not show significant correlations with any BFI dimensions. It is concluded that the new scale format (QoE-2B) is promising in mitigating social desirability bias, offering a potentially more reliable assessment of empathy and raising questions about the different aspects of empathy each format may measure.

*Keywords:* social desirability bias, empathy, self-report measures, psychological assessment, personality.

## Sumário

Introdução.....	7
Objetivos.....	13
Método.....	14
Participantes.....	14
Instrumentos.....	14
Procedimentos.....	17
Riscos.....	18
Benefícios.....	19
Resultados.....	19
Discussão.....	24
Considerações Finais.....	25
Referências.....	26

## **Introdução**

A avaliação psicológica (AP), área da psicologia responsável por medir processos psíquicos, ganhou destaque como campo de estudos nas últimas décadas. Isso é evidenciado pelo aumento significativo no número de artigos científicos publicados a partir da década de 2000 (Bueno & Peixoto, 2018). Nesse contexto, o Conselho Federal de Psicologia (CFP, 2022) define a avaliação psicológica como um processo estruturado de investigação de fenômenos psicológicos com o objetivo de auxiliar na tomada de decisão. Sob essa perspectiva, Hutz et al. (2015) enfatizam que a AP pode ir além da aplicação de testes, envolvendo uma gama diversificada de instrumentos e métodos como entrevistas, observação clínica, dinâmicas de grupo e questionários. Nesse sentido, em síntese, de acordo com Cohen, Swerdlik e Sturman (2014), a avaliação psicológica compreende a coleta e integração de dados psicológicos para realizar estimativas sobre características individuais, exigindo habilidades e treinamento específicos.

A AP está submetida a normativas e regras que validam a usabilidade de seus instrumentos e métodos, como, por exemplo, American Educational Research Association nos Estados Unidos, ou a British Psychological Society no Reino Unido. No Brasil, esse papel cabe ao Sistema de Avaliação de Testes Psicológicos (SATEPSI), que é responsável por avaliar a qualidade técnico-científica de instrumentos psicológicos para uso profissional (Conselho Federal de Psicologia, Resolução nº 31, art. 1º, 2022, p. 1). Dessa maneira, a responsabilidade e a segurança da avaliação psicológica passam por um crivo científico e ético para serem implementadas. Segundo Bueno e Peixoto (2018), fica clara a responsabilidade do psicólogo que irá escolher um instrumento para compor um processo de avaliação psicológica, uma vez que será necessário acessar as evidências de validade frente o

grupo de interesse e ao contexto onde o instrumento será utilizado, para que se possa criar bases de sustentação e, portanto, validar as interpretações psicológicas dos escores observados na testagem.

Entretanto, apesar de esforços e regulamentações, a AP ainda enfrenta desafios na obtenção de dados precisos. Dentre os potenciais fatores geradores de erros de medida, destaca-se a desejabilidade social (DS), que, de acordo com Almiro (2017), consiste no enviesamento de respostas, podendo ser entendida como uma tendência dos sujeitos a atribuírem para si próprios atitudes ou comportamentos socialmente desejáveis e, também, para tirarem de si mesmos a presença de atitudes ou comportamentos de valores socialmente indesejáveis. Logo, a desejabilidade social pode ser entendida como um obstáculo em estudos que envolvam o autorrelato de participantes. Isso porque o autorrelato é caracterizado pela pessoa de interesse ser a própria fonte de informação do pesquisador, então a forma em que as opções de resposta são apresentadas é de fundamental importância (Lima-Costa & Bonfá-Araújo, 2022).

A DS pode impactar pesquisas de forma significativa, alterando dados e colocando em questão o resultado da coleta de dados. Por exemplo, ao responder a um questionário de emprego que utiliza escala Likert para avaliar cooperação, com questões como: “Eu sou uma pessoa cooperativa”, é provável que as pessoas marquem opções próximas do “sim”, independentemente de serem cooperativas de fato, pois há uma expectativa social de que candidatos cooperativos sejam mais bem vistos e, portanto, preferidos para a vaga. Segundo Paulhus (1984) a desejabilidade social pode ocorrer de forma intencional ou não, sendo tais respostas indevidas entendidas de duas formas: autoengano e gerenciamento de impressão. No autoengano, as respostas imprecisas seriam motivadas por características de personalidade inflada e autoestima grande, onde o indivíduo acredita sinceramente em suas

autopercepções positivas. Em contrapartida, o gerenciamento de impressão ocorre quando o participante responde intencionalmente para causar uma boa impressão de si.

Pesquisas recentes sobre a desejabilidade social revelam o seu impacto nos resultados de estudos com autorrelato, além de possibilidades de como minimizar seus efeitos. Pode-se exemplificar a pesquisa de Ternovski e Orr (2022) que destacaram o aumento significativo no número de participantes desatentos em pesquisas online, que fornecem dados demográficos menos confiáveis, respostas menos estáveis e sistematicamente diferentes dos participantes atentos. Além disso, Bischof et al. (2024) argumentaram que as estratégias para controlar a DS variam de acordo com os objetivos da pesquisa. Os pesquisadores sugeriram três objetivos para avaliar a desejabilidade social: 1 assegurar que a desejabilidade social, e não fatores de confusão, seja medida nas respostas; 2 verificar se a manipulação da desejabilidade social foi efetiva; 3 utilizar diversas medidas para avaliar a influência da desejabilidade social. No entanto, os pesquisadores reconhecem as dificuldades de implementar os três objetivos, por isso, recomendam priorizar pelo menos dois desses objetivos, dependendo das especificidades de cada estudo. Nesse sentido, o presente estudo seguirá as diretrizes propostas pelos autores, priorizando os objetivos 1 e 3, que serão detalhados nas seções subsequentes.

Considerando a importância de aferir a desejabilidade social e a sua influência nas respostas, a escolha do instrumento para mensurar as respostas torna-se crucial. Nesse contexto, Pasquali (1998) afirma que é imprescindível considerar as características, limitações e fundamentação teórica do instrumento a ser utilizado. Uma quantidade bastante alta de testes de autorrelato utilizam escalas tipo Likert, criada por Rensis Likert (1932) para mensurar atitudes no contexto das ciências comportamentais a partir de categorias ordenadas para respostas. De acordo com Costa Jr. et al. (2024) a escala Likert pode medir as seguintes variações: Concordância, frequência, importância e probabilidade, sendo que tais variações

são destrinchadas em itens gradativos do mínimo ao máximo. No entanto, Costa Jr, et al. (2024) também apontam as limitações da escala em capturar a diversidade de respostas dos participantes, sugerindo a criação de variações como uma forma de complementar sua utilização. Além disso, pesquisas indicam que os níveis de classificação dos itens nas escalas Likert estarão relacionados a preocupações com a desejabilidade social e essa tendência em alguns avaliadores pode produzir um fator que está positivamente relacionado as escalas do Modelo de Cinco Fatores (FFM) (Bäckström & Björklund, 2013).

Ao considerar que Hutz et al. (2015) defendem que a AP visa gerar hipóteses e diagnósticos sobre diversos aspectos do indivíduo, como o funcionamento intelectual e a personalidade, a escala tipo Likert utilizada deve conter opções que vão ao encontro desses objetivos. Além disso, Wakita et al. (2012) afirmam ser fundamental que as escalas contenham distâncias psicológicas entre as opções, e que essas distâncias sejam iguais ou mais próximas possíveis, uma vez que, se essas distâncias entre as categorias são iguais, a escala tem medidas mais exatas do traço psicológico do entrevistado. Nesse estudo, os autores verificaram que o número de categorias na escala influencia essa distância psicológica, considerando as de 7 pontos mais propensas a distorções, especialmente em questões com conotações sociais negativas.

A mensuração da desejabilidade social pode ser realizada por meio de instrumentos específicos, como a Marlowe-Crowne Social Desirability Scale (EDSMC). Conforme explicado por Almiro (2017), esses instrumentos contêm itens que avaliam a tendência dos indivíduos a apresentar uma imagem socialmente desejável. A versão brasileira da EDSMC, adaptada por Gouveia et al (2009), intitulada de EDSMC-20, é composta por 20 itens nos quais os participantes indicam verdadeiro ou falso para situações sociais. Os itens da escala descrevem comportamentos socialmente desejáveis, mas pouco frequentes na realidade, como ‘Sou sempre educado(a), mesmo com pessoas desagradáveis’, ou comportamentos

indesejáveis, porém mais comuns, como ‘Gosto de fazer fofoca’. Quanto maior a pontuação, maior a desejabilidade social do respondente.

Sob a mesma perspectiva, é pertinente questionar a influência da desejabilidade social em construtos que avaliem comportamentos pró-sociais, como a empatia, afinal, os vieses de desejabilidade social são especialmente propensos a contaminar as medidas de empatia de autorrelato (Lovett & Sheffield, 2007). Assim, é fundamental considerar que a desejabilidade social pode influenciar a forma como os participantes respondem a itens relacionados à personalidade (Bäckström & Björklund, 2013). Já Decety e Cowell (2014) destacam que a empatia é um componente fundamental para as relações sociais, sendo ela tanto capaz de promover comportamentos pró-sociais, quanto gerar vieses. À vista disso, alguns estudos relacionam a empatia com componentes da personalidade, como explicado por Miguel et al. (2018), que demonstra uma relação significativa entre empatia e o fator socialização da Bateria Fatorial de Personalidade (BFP).

A empatia, definida como a capacidade de compreender e compartilhar as emoções dos outros (Decety & Cowell, 2014; Eres et al., 2015; Falcone et al., 2008), é um aspecto fundamental das interações sociais e, por isso, sua mensuração tem sido objeto de diversas pesquisas. Embora existam vários instrumentos confiáveis e válidos, ainda há espaço para melhorias, como apontado por Lima & Osório (2021) que explicam a necessidade de avanços em termos conceituais e teóricos para possibilitar o estabelecimento de modelos mais robustos, e salientam a importância de instrumentos com formatos alternativos para minimizar o viés de desejabilidade social em instrumentos de autorrelato.

Dada essa necessidade de minimizar o viés de desejabilidade social, é surpreendente que, apesar dos grandes avanços recentes das pesquisas sobre a DS, de acordo com um levantamento inicial, não há estudos analisando a relação entre desejabilidade social e as

escalas Likert em testes de empatia de autorrelato, assim como relacionando com traços de personalidade. Portanto, a presente pesquisa pretende preencher essa lacuna, explorando uma possível relação entre a DS e a autorreflexão nos testes de autorrelato, de forma a ir ao encontro dos pressupostos presentes na literatura (Pasquali, 1998; Wakita et al., 2012; Bischof et al., 2024; Almiro, 2017). Propõe-se, assim, a aplicação de testes de empatia de autorrelato em conjunto com escalas de desejabilidade social e de fatores de personalidade.

Ademais, embora os estudos relacionados a DS (Costa & Hauck Filho, 2017; Bäckström & Björklund, 2013) tenham se concentrado na influência nas respostas, eles não investigaram se as próprias opções da escala Likert poderiam contribuir para esse viés. Com o objetivo de responder a essa questão ainda não explorada, este trabalho propôs a adaptação da escala Likert em testes de empatia de autoaplicação, substituindo as opções de resposta (tradicionalmente apresentadas em um gradiente de intensidade) por alternativas que demonstram a frequência de certos comportamentos e incentivem uma autorreflexão mais aprofundada sobre os comportamentos avaliados nesses instrumentos. Desse modo, além de investigar se há relação entre autorrelato de empatia e DS, buscou-se verificar se esse novo formato também apresenta relação com DS.

Para investigar essa questão, esta pesquisa utilizou o Questionário Online de Empatia (QoE), um dos testes que compõem a Bateria Online de Inteligência Emocional (BOLIE) aprovada para utilização profissional pelo SATEPSI. O QoE é um teste autoaplicável que avalia a empatia por meio de três fatores: empatia afetiva, cognitiva e compassiva (Miguel, 2021), em uma ampla faixa etária de 9 a 88 anos. Sua relevância para esta pesquisa se dá não apenas pela capacidade de mensurar a empatia, mas também por permitir a investigação de como esse construto, frequentemente associado a comportamento socialmente desejáveis, é retratado em escalas de autorrelato, especialmente no que tange a desejabilidade social.

Com base na literatura que aponta a significativa influência da desejabilidade social nos testes de empatia de autorrelato, o presente estudo teve como objetivo investigar se um formato adaptado de escala (QoE-2B), focado na frequência de comportamentos empáticos, seria capaz de mitigar a tendência dos participantes de apresentar uma autoimagem idealizada ou socialmente aceitável, em comparação com o formato tradicional (QoE-2S). Para tanto, foram examinadas as relações das duas versões da escala de empatia com a desejabilidade social e com os traços de personalidade, avaliados por uma medida dos cinco grandes fatores. Este relatório detalhará a metodologia empregada, os resultados obtidos e as discussões sobre suas implicações para a teoria e a prática da avaliação da empatia."

### **Objetivos**

O presente estudo teve como objetivos principais: a) avaliar a relação entre a desejabilidade social e os resultados de testes de empatia autoadministrados que utilizam escala Likert; b) verificar se o novo formato de escala de resposta , focado na frequência de ocorrência de comportamentos empáticos, mitigou a influência da desejabilidade social nas respostas dos participantes, em comparação com a escala de resposta tradicional ; e c) investigar as relações entre a desejabilidade social, a empatia (em seus diferentes formatos de medida) e os traços de personalidade. Espera-se que os achados deste estudo contribuam significativamente para a construção do conhecimento acerca da desejabilidade social em testes de autorrelato e para o aprimoramento de instrumentos de avaliação da empatia.

## Método

### Participantes

A amostra deste estudo foi composta por 1.005 participantes, dos quais 241 se identificaram como do sexo masculino, 747 como do sexo feminino e 17 não declararam o sexo. As idades variaram entre 18 e 66 anos ( $M = 33,61$ ;  $DP = 10,65$ ). Os participantes foram recrutados por meio de estratégias de divulgação em redes sociais (Facebook, Instagram, WhatsApp) e por meios físicos, como panfletos e cartazes afixados em instituições de ensino superior. Os participantes foram divididos em 3 grupos:

- Grupo 1 ( $n=340$ ): Respondeu a versão QoE-2A, composta por 30 itens do QoE selecionados do QoE adaptativo, apresentados com a escala Likert original do teste.
- Grupo 2 ( $n=384$ ): Respondeu a versão QoE-2B, composta por 30 itens selecionados do QoE adaptativo (os mesmos do QoE-2A), apresentados com a escala Likert modificada para reduzir o viés de desejabilidade social.
- Grupo 3 ( $n=281$ ): Respondeu ao QoE-2A (escala Likert tradicional) e o QoE-2B (escala Likert alterada).

### Instrumentos

*Questionário Online de Empatia (QoE)*: A versão original é composta por 23 itens. Entretanto, há pesquisas para ampliar o banco de itens do QoE, que atualmente em sua versão adaptativa possui 113 itens, distribuídos em 3 fatores: 1 Responsividade, relacionado à empatia afetiva, que consiste em experimentar as mesmas emoções que as outras pessoas estão vivenciando; 2 Respeito, relacionado à empatia cognitiva, ou seja, compreender pensamentos, sentimentos e reações emocionais das outras pessoas; 3 Disponibilidade, relacionada à empatia passiva, isto é, desejo de se disponibilizar para ajudar as outras

pessoas. Esses fatores estão presentes em frases assinaladas em escala Likert de cinco pontos (1 Nunca; 2 Poucas vezes; 3 Às vezes sim, às vezes não; 4 Muitas vezes; 5 Sempre). (Miguel, 2021).

As pesquisas com QoE continuam em seu laboratório, sendo que atualmente dispõe-se de um banco de 113 itens para avaliar o construto, e não apenas os 23 itens da versão original. Assim, para a presente pesquisa, selecionou 30 itens da versão estendida do QoE, que foram distribuídos igualmente entre os três fatores (empatia afetiva, cognitiva e compassiva). Cada fator foi avaliado por meio de 5 itens positivos e 5 negativos, com frases como “Dedico parte do meu tempo para entender as outras pessoas” e “Eu faço trabalhos voluntários”. Essa nova versão do QoE (QoE-2) foi estudada em 3 grupos: o primeiro grupo com a escala Likert de 5 pontos original, que varia de ‘Nunca’ a ‘Sempre’ (Versão QoE-2A), e o grupo com a nova escala de 5 pontos que varia de ‘1 Deveria acontecer muito menos’ a ‘5 Deveria acontecer muito mais’ (Versão QoE-2B). O terceiro grupo respondeu as duas versões anteriores. É importante ressaltar que ambos os grupos responderam às mesmas 30 perguntas, sendo apenas as escalas de resposta que diferiram entre si.

*Versão QoE-2A* composta por 30 itens distribuídos em três fatores: empatia afetiva, cognitiva e compassiva. Cada fator contendo 5 positivos e 5 negativos, dispostos em frases assinaladas em escala Likert tradicional de cinco pontos (1 Nunca; 2 Poucas vezes; 3 Às vezes sim, às vezes não; 4 Muitas vezes; 5 Sempre).

*Versão QoE-2B* Composta pelos mesmos 30 itens do QoE-2A, mas com escala Likert alterada de 5 pontos (1 Deveria acontecer muito menos; 2 Deveria acontecer menos; 3 Não precisa acontecer nem mais nem menos; 4 Deveria acontecer mais; 5 Deveria acontecer muito mais).

Além do QoE, a pesquisa também utilizou a versão brasileira da Marlowe-Crowne Social Desirability Scale, adaptada por Gouveia et al. (2009) para identificar traços de desejabilidade social dos entrevistados, e o Big Five Inventory 2 - Short Version (BFI-2-S), que mede os cinco grandes traços de personalidade e seus subfatores. Dessa forma, as respostas dos entrevistados ao QoE com a nova escala puderam ser correlacionadas com instrumentos já validados que avaliam personalidade e desejabilidade social.

*Escala de Desejabilidade Social de Marlowe-Crowne (EDSMC-20)*, versão brasileira de Gouveia et al. (2009): Escala de autorrelato que mede a tendência de um indivíduo a apresentar uma imagem socialmente desejável. Composta por 20 itens que descrevem tanto comportamentos socialmente aceitáveis, mas improváveis, quanto comportamentos indesejáveis, mas comuns. Os participantes indicam se concordam ou discordam dessas informações (verdadeiro ou falso). E a pontuação total indica o grau de desejabilidade social presente nas respostas, quanto maior a pontuação, maior a desejabilidade social do indivíduo.

*Big Five Inventory 2 - Short Version (BFI-2-S)*: O BFI-2-SV (Pires et al., 2023), adaptado da versão original de Soto e John (2017), é um instrumento de autoavaliação que quantifica os cinco grandes fatores da personalidade: Extroversão (6 itens), Agradabilidade/Amabilidade (6 itens), Conscienciosidade (6 itens) Neuroticismo (6 itens) e Abertura à Experiência (6 itens), e seus subfatores: Sociabilidade, Assertividade e Nível de Energia (Extroversão), Compaixão, Respeito e Confiança (Agradabilidade/Amabilidade), Organização, Produtividade e Responsabilidade (Conscienciosidade), Ansiedade, Depressão e Volatilidade Emocional (Neuroticismo) e Senso Estético, Curiosidade Intelectual e Imaginação/Criatividade (Abertura à Experiência), contendo 2 itens para cada subfator. A pontuação em cada fator e subfator é obtida pela soma das respostas aos itens correspondentes, numa escala Likert de 5 pontos. A escala total varia de 30 a 150 pontos, enquanto cada fator varia de 6 a 30 pontos e cada subfator de 2 a 10 pontos.

## **Procedimentos**

### *Etapa 1 - Condução de procedimentos éticos*

Após a aprovação do projeto Comitê de Ética em Pesquisa (CEP) UFSCar, os participantes foram convidados a participar desta pesquisa por meio de divulgação digital em redes sociais, e divulgação física como panfletos e cartazes.

### *Etapa 2 - Recrutamento dos participantes*

Apenas os indivíduos que consentiram formalmente com o Termo de Consentimento Livre e Esclarecido (TCLE), alocado em um ambiente online, de domínio próprio do orientador, programados em linguagem ASP.NET. Seguindo as diretrizes nacionais e internacionais para aplicação em ambiente informatizado (Conselho Federal de Psicologia, 2022; International Test Commission, 2005), o Termo de Consentimento Livre e Esclarecido (TCLE) era apresentado na tela, em vez de em papel, sendo a primeira e única tela vista pelo participante que acessa a pesquisa no website. Caso concordasse com a participação, o indivíduo pressionava o botão “Concordo em participar”. Ademais, fora assegurado o sigilo das informações obtidas, a forma de divulgação dos dados e a possibilidade de desistência do participante a qualquer momento da coleta, sem nenhum tipo de prejuízo.

### *Etapa 3 - Coleta de dados*

Após a assinatura do TCLE, disponibilizado no ambiente online, eram apresentados os seguintes testes e escala: QoE-2A ou QoE2-B, BFI, EDSMC-20. Os dados dos participantes estão armazenados em um banco de dados eletrônico, protegido por senha conhecida apenas pelo pesquisador responsável. Não há identificação do participante, como

nome ou CPF, apenas dados sociodemográficos (idade, sexo, escolaridade e estado onde reside).

#### *Etapa 4 - Análise de dados*

Para se atingir os objetivos propostos pela pesquisa, executaram-se as seguintes análises estatísticas dos dados obtidos:

- Teste ANOVA (ou medida compatível, a depender da normalidade dos dados) entre Grupo 1, Grupo 2 e Grupo 3, para verificar se há diferenças na distribuição das pontuações do QoE dependendo da modalidade de escala.
- Correlações entre QoE-2A com DS, assim como correlações entre QoE-2B com DS, a fim de verificar se uma das versões está mais suscetível à desejabilidade social.
- Correlações entre QoE-2A com BFI-2-S, assim como correlações entre QoE-2B com BFI-2-S, a fim de verificar se uma das versões demonstra melhores evidências de validade.
- No Grupo 3, criação de um escore que representa a diferença entre a pontuação de QoE-2A e QoE-2B. Então, correlação desse escore com DS e BFI-2-S, a fim de verificar se o distanciamento entre as escalas está associado a maior desejabilidade social e/ou traços de personalidade.

#### **Riscos**

A participação nesta pesquisa apresentou riscos leves, esses incluíram a possibilidade de: cansaço, inibição para participar do estudo, sentir-se entediado durante o procedimento, riscos relacionados ao uso de meios virtuais ou alguma reação emocional às perguntas apresentadas.

Em relação à plataforma digital, os riscos possíveis foram falhas na conexão que podem inviabilizar a aplicação do teste. Como qualquer plataforma digital, há risco de violação por terceiros. Nesse sentido, como já mencionado, assegura-se que dados pessoais que possam identificar o participante não serão coletados e/ou armazenados.

### **Benefícios**

O estudo poderá auxiliar no desenvolvimento de instrumentos mais robustos e adaptados para avaliar a empatia, oferecendo opções metodológicas que poderão reduzir o viés de desejabilidade social em pesquisas futuras.

A pesquisa permitiu a comparação entre as versões das escalas (QoE-2A e QoE-2B), contribuindo para identificar medidas mais fidedignas no campo da psicologia.

Os resultados possibilitam conhecimento sobre os traços de empatia e suas relações com personalidade e desejabilidade social na população brasileira, possibilitando a ampliação das possibilidades de intervenções e estudos na área.

A participação no estudo ofereceu aos indivíduos uma oportunidade para refletir sobre suas próprias percepções e comportamentos relacionados à empatia e à personalidade.

### **Resultados**

#### **Tabela 1**

*Estatística descritiva dos participantes - Sexo*

Sexo	Número de Participantes	Percentual (%)
Masculino	241	23.98
Feminino	747	74.33
Não Informado	17	1.69
Total	1005	100

A Tabela 1 apresenta a distribuição de gênero da amostra. Observa-se uma predominância de participantes do sexo feminino (74.33%), enquanto os participantes do sexo masculino correspondem a 23.98%.

## **Tabela 2**

### *Estatística descritiva dos participantes - escolaridade*

Escolaridade	Número de Participantes	Percentual (%)
Ensino Fundamental	6	.6
Ensino Médio	106	10.55
Ensino Superior	571	56.82
Pós-Graduação	253	25.17
Não Informado	69	6.87
Total	1005	100

Conforme exposto na Tabela 2, a amostra do estudo apresenta um alto nível de escolaridade. A maioria dos participantes possuía ensino superior (56.82%) ou pós graduação (25.17%), totalizando 81,99%. Uma parcela menor possuía ensino médio (10.55%) ou fundamental (.6%).

**Tabela 3***Estatística descritiva dos participantes - Região*

Região	Frequency	Percent
Norte	33	33.28
Nordeste	165	16.42
Centro-Oeste	64	6.37
Sudeste	580	57.71
Sul	91	45786
Não Informado	72	7.16
Total	1005	100

A tabela 3 detalha a distribuição geográfica dos participantes da pesquisa. A região sudeste do Brasil foi a mais representada na amostra, compreendendo 57.71% do total. As demais regiões contribuíram com parcelas menores de participantes.

**Tabela 4**

*Tabela de Confiabilidade (Alpha de Cronbach) da Escala de Desejabilidade Social EDSMC-20*

Número de itens	Alfa de Cronbach (original)	Alfa de Cronbach (Final, após correção)
20	0.56	0.69

A Tabela 4 apresenta a confiabilidade interna da Escala de Desejabilidade Social (EDSMC-20), avaliada por meio do coeficiente Alpha de Cronbach. Inicialmente, a escala obteve um Alpha de 0.56. Após a identificação e correção de dois itens com pontuação invertida, o Alpha de Cronbach da escala foi aprimorado para 0,69, sendo um valor aceitável para o estudo.

### Tabela 5

*Correlação de Pearson entre QoE e EDSMC-20 Corrigida*

Variável QoE	EDSMC-20 Corrigida
<b>QoE-2A</b>	.35**
1 Responsividade	.16**
2 Respeito	.37**
Testes	.37**
<b>QoE-2B</b>	-.01
1 ResponsividadeQ	-.02
2 Respeito	-.04
Disponibilidade	.04

Nota: \*\*p < 0,01

A tabela 5 eibe as correlações de Pearson entre as duas versões da escala de empatia (QoE-2A e QoE-2B) e a escala de Desejabilidade Social (EDSMC-20 corrigida). O QoE-2A apresenta correlação positiva e estatisticamente significativa com a desejabilidade social ( $r=$

.35). Em contraste, a versão QoE-2B não apresentou correlação significativa com a desejabilidade social ( $r = -.01$ ).

**Tabela 6**

*Correlação de Pearson entre QoE e BFI-2-S*

	BFI - Extroversão	BFI - Agradabilidade	BFI - Conscienciosidade	BFI - Neuroticismo	BFI - Abertura à Experiência
QoE-2A	.20**	.67**	.20**	-.21**	.22**
1 Responsividade	.21**	.49**	.07	-.05	.16**
2 Respeito	0.04	.55**	.19**	-.28**	.23**
3 Disponibilidade	.23**	.63**	.23**	-.22**	.17**
QoE-2B	.03	-.05	.04	-.02	-.07
1 Responsividade	-.01	-.04	.03	-.12**	-.05
2 Respeito	.06	-.08	.04	.06	-.06
3 Disponibilidade	.01	-.01	.03	0	-.06

*Nota:* A correlação é estatisticamente significativa no nível de 0,01 (bicaudal)

A tabela 06 detalha as correlações de Pearson entre as escalas de empatia (QoE-2A e QoE-2b) e as dimensões do inventário dos Cinco grandes Fatores de Personalidade (BFI-2-S). A QoE-2A apresentou correlação significativas com diversas dimensões do BFI. Em contrapartida, a QoE-2B não demonstrou correlações significativas com nenhuma das dimensões do BFI.

## Discussão

A presente pesquisa teve como objetivo principal examinar o impacto da desejabilidade na avaliação da empatia por meio de tese de autorrelato e, simultaneamente, explorar a eficácia de um novo formato de escala (QoE-2B) na minimização desse viés. Os resultados obtidos para a versão tradicional da escala, a QoE-2A, confirmam a hipótese inicial de vulnerabilidade. Evidenciou-se uma correlação positiva e estatisticamente significativa entre a QoE-2A e a Escala de Desejabilidade Social (EDSMC-20), reforçando o consenso na literatura de que a tendência a responder de forma socialmente aceitável pode influenciar as pontuações em medidas de autorrelato de construtos socialmente valorizados, como empatia.

O achado mais promissor do estudo reside na performance do novo formato da escala, a QoE-2B, que foca na frequência de comportamentos empáticos. Diferentemente da QoE-2A, a QoE-2B não demonstrou correlação significativa com a desejabilidade social. Este resultado sugere que a estratégia de desenvolver uma escala que avalia a empatia por meio indicadores comportamentais mais objetivos pode ser eficaz na mitigação da influência desse viés de resposta. Essa descoberta possui implicações significativas para a pesquisa em empatia, oferecendo um caminho para o desenvolvimento de instrumentos menos suscetíveis a distorções, o que pode levar a uma avaliação mais precisa do construto em diversos contextos.

Adicionalmente, as análises de validade convergente, por meio das correlações com o Big Five Inventory (BFI-2-S), ofereceram insights complementares sobre as duas versões da escala. A QoE-2A apresentou correlações esperadas e significativas como os fatores de personalidade, como irritabilidade, o que corrobora sua validade e alinhamento com modelos teóricos que relacionam empatia a traços de personalidade. Em contraste, a QoE-2B não apresentou correlações significativas com nenhum dos fatos do BFI. Esse padrão sugere que a QoE-2B pode estar mensurando um aspecto da empatia distinto daquele capturado pelas

medidas de traço de personalidade, ou capturando traços e valores psicológicos ainda desconhecidos, o que merece aprofundamento em investigações futuras.

Em síntese, o estudo contribui para o campo da psicometria ao validar a vulnerabilidade de escalas de autorrelato de empatia à desejabilidade social e, mais importante, ao propor e testar um formato de escala alternativo com potencial para mitigar esse viés. A demonstração de que a QoE-2B não se correlaciona com a desejabilidade social representa um avanço metodológico significativo. Contudo, é importante reconhecer as limitações da pesquisa, incluído a utilização de uma amostra com conveniência com predominância feminina, alto nível educacional e concentração geográfica, o que pode restringir a generalização dos resultados.

Considerando as descobertas e limitações, pesquisas futuras são imperativas para consolidar e expandir esses achados. Recomenda-se a realização de análises fatoriais com outros instrumentos, a fim de investigar a correlação da escala de resposta com outros instrumentos. A reaplicação do estudo em amostras mais diversas e em diferentes contextos cruciais para fortalecer a robustez e a aplicabilidade da nova escala.

### **Considerações Finais**

As considerações finais desta pesquisa reforçam os principais achados e suas implicações no campo da psicometria e da avaliação da empatia. O estudo demonstrou de forma clara a vulnerabilidade da versão tradicional da escala de empatia (QoE-2A) ao invés da desejabilidade social, confirmando a necessidade de abordagens alternativas. O avanço mais significativo reside na avaliação de que a escala de resposta adaptada (QoE-2B), com foco na frequência de comportamentos, efetivamente minimiza essa influência. Apresentando-se como uma ferramenta promissora para uma avaliação mais fidedigna da empatia.

Esta pesquisa oferece uma contribuição valiosa ao abordar criticamente a desejabilidade social na avaliação da empatia e ao propor uma solução inovadora com a QoE-2B. No entanto, é fundamental reconhecer as limitações do estudo, particularmente em relação à composição da amostra de conveniência, predominantemente feminina e com alta escolaridade, que pode restringir a generalização dos achados.

## Referências

- Almiro, P. A. (2017). Editorial. *Revista Avaliação Psicológica*, 16(03).  
<https://doi.org/10.15689/ap.2017.1603.ed>
- BÄCKSTRÖM, M., & BJÖRKLUND, F. (2012). Social desirability in personality inventories: Symptoms, diagnosis and prescribed cure. *Scandinavian Journal of Psychology*, 54(2), 152–159. <https://doi.org/10.1111/sjop.12015>
- Bischof, D., Tim Lars Allinger, Le, M., & Vrede, K. (2024). *A Comprehensive Test of the Most Promising Method to Capture Social Desirability Bias in Online Surveys*.  
<https://doi.org/10.31219/osf.io/ew95m>
- Bispo Júnior, J. P. (2022). Social desirability bias in qualitative health research. *Revista de Saúde Pública*, 56(56), 101. <https://doi.org/10.11606/s1518-8787.2022056004164>
- Bueno, J. M. H., & Peixoto, E. M. (2018). Avaliação Psicológica no Brasil e no Mundo. *Psicologia: Ciência E Profissão*, 38(spe), 108–121.  
<https://doi.org/10.1590/1982-3703000208878>
- Conselho Federal de Psicologia [CFP]. (2022). *Cartilha Avaliação Psicológica 2022*. CFP.  
<https://site.cfp.org.br/publicacao/cartilha-avaliacao-psicologica-2022/>
- COSTA JÚNIOR, J. F., CABRAL, E. L., SOUZA, R. C., BEZERRA, D. M. C., & SILVA, P. T. de F. (2024). Um estudo sobre o uso da escala de Likert na coleta de dados

- qualitativos e sua correlação com as ferramentas estatísticas. *Contribuciones a Las Ciencias Sociales*, 17(1), 360–376. <https://doi.org/10.55905/revconv.17n.1-021>
- Decety, J., & Cowell, J. M. (2014). The complex relation between morality and empathy. *Trends in Cognitive Sciences*, 18(7), 337–339. <https://doi.org/10.1016/j.tics.2014.04.008>
- Eres, R., Decety, J., Louis, W. R., & Molenberghs, P. (2015). Individual differences in local gray matter density are associated with differences in affective and cognitive empathy. *NeuroImage*, 117, 305–310. <https://doi.org/10.1016/j.neuroimage.2015.05.038>
- Falcone, E. M. de O., Ferreira, M. C., Luz, R. C. M., Fernandes, C. S., Faria, C. de A., D'Augustin, J. F., ...Pinho, V. D. (2008). Inventário de Empatia (I.E.): Desenvolvimento e validação de uma medida brasileira. *Avaliação Psicológica*, 7(2), 321-334.
- Gouveia, V. V., Guerra, V. M., Sousa, D. M. F. de, Santos, W. S., & Costa, J. de M. (2009). Escala de Desejabilidade Social de Marlowe-Crowne: evidências de sua validade fatorial e consistência interna. *Avaliação Psicológica*, 8(1), 87–98.
- Hutz, C. S., Bandeira, D. R., & Trentini, C. M. (2015). *Psicometria*. Artmed.
- Jamieson, S. (2004). Likert scales: how to (ab)use them. *Medical Education*, 38(12), 1217–1218. <https://doi.org/10.1111/j.1365-2929.2004.02012.x>
- Larson, R. B. (2019). Controlling social desirability bias. *International Journal of Market Research*, 61(5), 534–547. <https://doi.org/10.1177/1470785318805305>
- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, 140, 1–55.
- Lima Costa, A. R., & Hauck Filho, N. (2017). Menos desejabilidade social é mais desejável: Neutralização de instrumentos avaliativos de personalidade. *Interação Em Psicologia*, 21(3). <https://doi.org/10.5380/psi.v21i3.53054>

- Lima, F. F. de, & Osório, F. de L. (2021). Empathy: Assessment Instruments and Psychometric Quality – A Systematic Literature Review With a Meta-Analysis of the Past Ten Years. *Frontiers in Psychology*, 12. <https://doi.org/10.3389/fpsyg.2021.781346>
- Lima, F. F. de, & Osório, F. de L. (2021). Empathy: Assessment Instruments and Psychometric Quality – A Systematic Literature Review With a Meta-Analysis of the Past Ten Years. *Frontiers in Psychology*, 12. <https://doi.org/10.3389/fpsyg.2021.781346>
- Lima-Costa, A. R., & Bonfá-Araujo, B. (2022). Construção de escalas de autoinforme: ¿Qué hacer? *Avaliação Psicológica*, 21(3), 329–338. <https://doi.org/10.15689/ap.2022.2103.21860.09>
- LOVETT, B., & SHEFFIELD, R. (2007). Affective empathy deficits in aggressive children and adolescents: A critical review. *Clinical Psychology Review*, 27(1), 1–13. <https://doi.org/10.1016/j.cpr.2006.03.003>
- Meisters, J., Hoffmann, A., & Musch, J. (2020). Controlling social desirability bias: An experimental investigation of the extended crosswise model. *PLOS ONE*, 15(12), e0243384. <https://doi.org/10.1371/journal.pone.0243384>
- Miguel, F. K. (2021). *Bolie - Bateria online de inteligência emocional* (1st ed.). Vetor Editora.
- Miguel, F. K., Hashimoto, E. S., Gonçalves, E. R. D. S., Oliveira, G. T., & Wiltenburg, T. D. (2018). Estudos de validade do questionário online de empatia. *Temas Em Psicologia*, 26(4), 2203–2216. <https://doi.org/10.9788/tp2018.4-18pt>
- Norman, G. (2010). Likert scales, levels of measurement and the “laws” of statistics. *Advances in Health Sciences Education : Theory and Practice*, 15(5), 625–632. <https://doi.org/10.1007/s10459-010-9222-y>

- Nunes, C. H. S. S., & Hutz, C. S. (2005). O modelo dos cinco grandes fatores de personalidade. Em R. Primi (Org.), *Temas em avaliação psicológica* (p. 87–106). Casa do Psicólogo
- Pasquali, L. (1998). Princípios de elaboração de escalas psicológicas. *Revista de Psiquiatria Clínica, 25*(5), 206–213.
- Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology, 46*(3), 598–609.  
<https://doi.org/10.1037/0022-3514.46.3.598>
- Resolução nº 031, de 15 de dezembro de 2022.* (2022). CFP. Brasília, DF
- Ronald Jay Cohen, Swerdlik, M. E., & Sturman, E. D. (2014). *Testagem e Avaliação Psicológica - 8.ed.* AMGH Editora.
- Rosenman, R., Tennekoon, V., & Hill, L. G. (2011). Measuring bias in self-reported data. *International Journal of Behavioural and Healthcare Research, 2*(4), 320.  
<https://doi.org/10.1504/ijbhr.2011.043414>
- Ternovski, J., Orr, L., Kalla, J., & Aronow, P. (2022). A Note on Increases in Inattentive Online Survey-Takers Since 2020. *Journal of Quantitative Description: Digital Media, 2*. <https://doi.org/10.51685/jqd.2022.002>
- Urbina, S. (2007). *Fundamentos da testagem psicológica.* (pp. 11–40). Porto Alegre Artmed.
- Wakita, T., Ueshima, N., & Noguchi, H. (2012). Psychological Distance Between Categories in the Likert Scale. *Educational and Psychological Measurement, 72*(4), 533–546.  
<https://doi.org/10.1177/0013164411431162>