



UNIVERSIDADE FEDERAL DE SÃO CARLOS  
CENTRO DE EDUCAÇÃO E CIÊNCIAS HUMANAS  
PROGRAMA DE PÓS-GRADUAÇÃO EM PSICOLOGIA

**DENISE APARECIDA PASSARELLI**

*Análise do Comportamento e Preconceito Racial: Uma História de Pesquisa*

São Carlos - SP

Março de 2026

**DENISE APARECIDA PASSARELLI**

Análise do Comportamento e Preconceito Racial: Uma História de Pesquisa

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Psicologia da Universidade Federal de São Carlos, como parte dos requisitos para obtenção do título de Doutora em Psicologia, sob orientação do Prof. Dr. Júlio César Coelho de Rose e coorientação de Dr. Bryan Roche.

---

Pesquisa financiada pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), no âmbito de bolsa de Doutorado (Processos FAPESP nº 2021/04105-8 e nº 2023/06402-5) e pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), por meio do Programa de Excelência Acadêmica (PROEX, processo Nº 88881.185837/2025-01).

---

## Folha de Aprovação

---

Defesa da Tese de Doutorado da candidata Denise Aparecida Passarelli realizada em 13 de março de 2026.

### Comissão Julgadora:

Prof. Dr. Júlio César Coelho de Rose (UFSCar)

Prof. Dr. Nassim Chamel Elias (UFSCar)

Prof. Dr. William Ferreira Perez (Instituto Par)

Prof. Dr. Vinicius Pereira de Souza (UPM)

Prof. Dr. João Henrique Almeida (UNESP)

### **Apoio Financeiro**

A discente recebeu financiamento principal por meio de bolsa de doutorado da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, Processos nº 2021/04105-8 e nº 2023/06402-5) e financiamento parcial da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES, Processo nº 88887.611134/2021-00).

Cabe destacar que o presente projeto de doutorado foi desenvolvido com o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), por meio do Programa de Excelência Acadêmica (PROEX), vinculado ao Programa de Pós-Graduação em Psicologia da Universidade Federal de São Carlos (PPGpsi-UFSCar), com financiamento do Programa CAPES-PROEX-AUXPE Nº 3596/2025 (Processo Nº 88881.185837/2025-01). A infraestrutura necessária para a realização da pesquisa foi fornecida pelo Instituto Nacional de Ciência e Tecnologia sobre Comportamento, Cognição e Ensino (INCT-ECCE), com financiamento do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Processos No 465686/2014-1, 409051/2024-1) e FAPESP (Processo No 2014-50909-8).

## Agradecimentos

Esta tese é resultado de um percurso construído em interação com múltiplos contextos, práticas e interlocuções que tornaram possível sua realização. A produção científica aqui apresentada não é fruto de um esforço isolado, mas de contingências acadêmicas, institucionais e sociais que sustentaram, desafiaram e orientaram o desenvolvimento deste trabalho.

Agradeço aos espaços de formação e pesquisa que possibilitaram o diálogo rigoroso entre teoria, método e dados, bem como às instituições que ofereceram as condições materiais e estruturais necessárias para a realização desta investigação. Reconheço, em especial, a importância de ambientes, não apenas acadêmicos, que valorizam o debate crítico, a autonomia intelectual e o compromisso ético com a ciência. Nesse sentido, começo agradecendo às instituições financiadoras FAPESP, CAPES e CNPq. Agradeço também ao Portal de Periódicos da CAPES, que viabilizou o acesso à literatura necessária para a construção desta tese: ao Laboratório de Estudos sobre Comportamento Humano (LECH), PPGPsi-UFSCar e ao Instituto Nacional de Ciência e Tecnologia sobre Comportamento, Cognição e Ensino (ECCE - INCT), que forneceram a infraestrutura para o desenvolvimento deste trabalho. Registro ainda um agradecimento especial ao Conselho Federal de Psicologia (CFP), ao Conselho Regional de Psicologia de São Paulo (CRP), à *Association for Behavior Analysis International* (ABAI) e ao *Experimental Analysis of Human Behavior – Special Interest Group* (EAHB SIG), pelas premiações que contribuíram para a visibilidade desta pesquisa.

Sou grata a todas as pessoas que, em diferentes momentos, contribuíram por meio de discussões, questionamentos, leituras, críticas e sugestões, ampliando o alcance e a consistência deste programa de pesquisa. Essas interlocuções foram fundamentais para o

refinamento conceitual e metodológico que sustenta a tese. Dada a quantidade de pessoas envolvidas e o risco de omissões, faço um agradecimento especial à tríade que contribuiu de forma incansável para a construção deste trabalho: Júlio de Rose, Táhcita Mizael e Bryan Roche. Adicionalmente, agradeço à comissão julgadora (Vinicius de Souza, Nassim Chanel Elias, William Perez e João de Almeida) pelas sugestões que possibilitaram melhorias significativas na qualidade desse trabalho.

Agradeço, ainda, com especial carinho, às participantes e aos participantes dos experimentos, cuja colaboração tornou possível a investigação empírica de fenômenos complexos relacionados ao preconceito racial e ao responder relacional. Por razões éticas e de sigilo, não menciono diretamente as escolas envolvidas, mas registro que sua abertura à pesquisa representa uma ação institucional concreta de enfrentamento ao racismo.

Por fim, e não menos importante, reconheço o apoio recebido fora do contexto acadêmico, nas relações cotidianas que ofereceram sustentação emocional, estabilidade e incentivo ao longo desse percurso. Esses apoios, muitas vezes silenciosos, foram essenciais para que este trabalho pudesse ser concluído. Não seria justo nomear pessoas específicas, pois não haveria espaço suficiente para fazer jus a todas; assim, este agradecimento se estende à minha família, ao meu companheiro de vida, à minha psicóloga e a todas as amigas, amigos e colegas que caminharam comigo ao longo desse processo.

*“Onde haja Fé, que eu leve a Dívida”*

*Marcondes Falcão Maia*

## SUMÁRIO

Resumo .....	15
Abstract.....	17
Prefácio.....	19
Capítulo 1. Introdução .....	25
Racismo e Viés Racial .....	26
Perspectivas da Análise do Comportamento sobre Atitudes, Preconceito, Preferências e Discriminação .....	30
A Caixa de Pandora do Preconceito Racial: Males Persistem, a Esperança Permanece ..	36
Capítulo 2. Análise do Comportamento, Racismo e Preconceito Racial: Onde estivemos e para onde vamos? .....	38
Background.....	38
Estudo 1 - Uma Revisão Sistemática de Dados Empíricos sobre Viés Racial Antinegro na Análise do Comportamento .....	41
Introdução .....	42
Método .....	49
Resultados .....	51
Características Metodológicas dos Estudos .....	53
Estudos Transversais e Correlacionais.....	60
Estudos de Medidas Comportamentais de Identificação de Preconceito Racial.....	64
Estudos de Intervenções no Viés Racial Antinegro .....	68
Discussão e Direções Futuras .....	75
Comentário.....	83
Capítulo 3. Medindo o Preconceito Racial.....	85
Background.....	85
Estudo 2 - Examinando o viés racial antinegro implícito e explícito em crianças: um delineamento de medidas repetidas utilizando o <i>Functional Acquisition Speed Test</i> (FAST) .....	92
Introdução .....	93
Método .....	101
Participantes e Local .....	101
Procedimento.....	102
Análise de Dados .....	106
Resultados.....	107

Escala de Valência Emocional .....	107
Function Acquisition Speed Test (FAST).....	109
Correspondência entre as Medidas.....	112
Discussão .....	114
Comentário.....	121
Estudo 3 - Identificação do preconceito racial em crianças com uso de diferentes medidas: Uma análise exploratória multimodal .....	125
Introdução .....	126
Método .....	131
Participantes e Local .....	131
Instrumentos .....	131
Teste da Boneca .....	133
Procedimento.....	135
Análise de Dados.....	139
Resultados .....	140
SAM .....	140
Teste da Boneca .....	148
FAST .....	149
Discussão .....	151
Considerações Finais do Capítulo.....	160
Capítulo 4. Intervenções voltadas para Redução do Preconceito Racial em Crianças.....	166
Estudo 4 - Reduzindo o Viés Racial por meio do Paradigma de Relações Conflitantes: Uma Replicação Sistemática de Mizael et al. (2016) .....	171
Introdução .....	172
Método .....	179
Participantes .....	179
Local, Equipamento e Estímulos.....	180
Procedimento.....	182
Medidas Avaliativas <i>Teste Relacional</i> .....	183
Escala Likert.....	185
Intervenção.....	189
Resultados .....	193
<i>Function Acquisition Speed Test (FAST)</i> .....	199

Discussão .....	203
Comentário: Replicar não é apenas repetir efeitos, mas identificar limites e gerar novas questões de pesquisa .....	215
Estudo 5 - Efeitos da Formação de Classes de Equivalência na Redução do Preconceito Racial em Crianças I: Evidências de Generalização e Manutenção .....	220
Introdução .....	221
Método .....	226
Participantes .....	226
Local e Estímulos .....	227
Instrumentos .....	227
Procedimento .....	229
Resultados .....	234
Matching-to-Sample .....	234
Teste Relacional .....	235
Escala de Valência Emocional ( <i>Likert</i> ) .....	238
Teste das Bonecas .....	239
Estudo 6 - Efeitos da Formação de Classes de Equivalência na Redução do Preconceito Racial em Crianças II: Evidências de Generalização, Manutenção e Validade Social ..	253
Introdução .....	254
Método .....	258
Intervenção .....	264
Validade Social .....	266
Resultados .....	267
Desempenho na Intervenção .....	267
Teste Relacional com Três Estímulos Comparações .....	268
Teste Relacional Preferências com Duas Comparações .....	272
Self-Assessment Manikin .....	275
Análise Integrativa dos Efeitos da Intervenção e Convergência entre Medidas .....	282
Validade Social .....	285
Discussão .....	286
Considerações Finais do Capítulo .....	297
Capítulo 5. Intervenção para Preconceito Racial em Adultos .....	302
Background .....	302

Estudo 7 - Mitigando o viés racial em adultos brancos por meio de treino relacional profilático com o Function Acquisition Speed Test (FAST).....	306
Materiais.....	311
Procedimento.....	318
Resultados.....	322
Desempenho nas Intervenções FAST .....	322
Dilema do Policial (PODT).....	325
Escala de Racismo Moderno Modificada (MMRS).....	332
Escala Likert.....	333
Correlações entre as Medidas.....	333
Discussão.....	334
Comentário Final do Capítulo.....	347
Considerações finais.....	351
Síntese 1 – Necessidade da sistematização de uma agenda de pesquisa sobre racismo e preconceito racial na Análise do Comportamento .....	351
Síntese 2 - Formulações teóricas não devem anteceder as observações experimentais .....	352
Síntese 4 – Preconceito racial como responder relacional sob múltiplos controles....	353
Síntese 5 – As medidas utilizadas mostraram estabilidade sob repetição espaçada ...	354
Síntese 6 – A idade da amostra mostrou-se uma variável funcionalmente relevante .	354
Síntese 7 – A formação de classes de equivalência é condição necessária para mudanças, mas não garante generalização.....	355
Síntese 8 – Resistência à mudança emerge cedo.....	356
Síntese 9 – Flexibilidade relacional envolve eliminar a raça como controle do responder .....	356
Síntese 10 – Validade social não é acessório, é a construção da relação da pesquisadora com todos os membros do ambiente da pesquisa.....	357
Síntese 11 – FAST pode ser usado como medida, mas também como ferramenta de treino.....	357
Síntese 12 – Reorganização relacional é funcionalmente distinta da formação de classes.....	359
Síntese 13 – A exposição às contingências que reforçam viés racial antinegro aprofunda o viés .....	360
Síntese 14 – Avançar exige precisão conceitual, controle experimental e humildade empírica.....	360

Síntese 15 –Intervir é manipular contingências, não corrigir pessoas .....	361
Referências .....	362
APÊNDICES .....	401
Anexo III - Instruções Fornecidas às Crianças por Instrumento .....	406

### Lista de Figuras

Figura 1. Diagrama PRISMA com procedimento de busca e resultados.....	53
Figura 1. Valência média atribuída a rostos negros e brancos nos diferentes subgrupos .....	110
Figura 1. Escala SAM apresentada aos participantes com exemplo de face apresentada.....	133
Figura 1. Desempenho Individual nas Sondas antes e Depois da Intervenção.....	270
Figura 1. Fluência de resposta (RRD) e diferencial de fluência de resposta (RFD) na condição de flexibilidade (C1) .....	322
Figura 2. Desempenho nos blocos consistente e inconsistente entre os subgrupos, destacando viés antinegro e variações no efeito FAST.....	113
Figura 2. Estímulos utilizados no Teste da Boneca Modificado.....	134
Figura 2. Exemplo dos dois tipos de tentativas apresentadas no Matching to Sample.....	184
Figura 2. Avaliações médias dos participantes antes e após a intervenção.....	238
Figura 2. Desempenho no Teste com duas Comparações Antes e Depois da Intervenção.....	273
Figura 2. Fluência de resposta ao longo de seis exposições aos blocos nas condições C2 e C3.....	324
Figura 3. Efeitos do gênero do estímulo facial e da raça da face sobre as avaliações de valência em crianças brancas e pardas.....	142
Figura 3. Escala Likert utilizada para avaliação da valência de faces e dos distratores.....	186
Figura 3. Exemplos de Tentativas apresentadas no Bloco Consistente e Inconsistente do FAST.....	188
Figura 3. Valência atribuída às faces negras e brancas no SAM.....	276

Figura 3. Acurácia das decisões em relação aos tipos de objeto para alvos negros e brancos, ao longo das condições experimentais.....	326
Figura 4. Efeitos do gênero do estímulo facial e da raça da face sobre as avaliações de <i>arousal</i> em crianças brancas e pardas.....	144
Figura 4. Visão Geral da Intervenção.....	192
Figura 4. Escores médios de critério de decisão (c) para alvos negros e brancos ao longo das quatro condições experimentais.....	329
Figura 5. Efeitos do gênero do estímulo facial e da raça da face sobre as avaliações de dominância em crianças brancas e pardas.....	147
Figura 5. Escores de sensibilidade (d') para alvos negros e brancos ao longo das quatro condições experimentais.....	331
Figura 5. Índice de Viés Racial Antinegro nas Linhas de Base Antes e Depois da Intervenção.....	195
Figura 6. Frequência de escolhas da boneca branca e negra em função do contexto da tarefa no Teste da Boneca.....	150
Figura 6. Viés Racial Médio Antes da Intervenção, depois da Intervenção e no seguimento de 15 dias após a intervenção.....	196
Figura 7. Inclinação Média nos Blocos Consistentes e Inconsistentes em participantes .....	151
Figura 7. Valência das faces negras e brancas antes e depois da intervenção.....	197
Figura 7. Diferença média nas avaliações em escala Likert entre faces negras e faces brancas.....	199
Figura 8. Desempenho nos Blocos Consistentes e Inconsistentes do FAST Antes e Após a Intervenção.....	200

### **Lista de Tabelas**

Tabela 1. Características Metodológicas dos Estudos Seleccionados.....	57
Tabela 1. Caracterização da Amostra.....	102
Tabela 1. Descrição dos Participantes e medidas que mostraram viés.....	180
Tabela 1. Caracterização dos participantes.....	226
Tabela 1. Moda das respostas no Teste da Boneca Adaptado por participante, atividade e tipo de resposta (pré- e pós-intervenção).....	278

Tabela 2. Média e desvio-padrão das avaliações em escala Likert de rostos negros e brancos obtidas ao longo das medidas repetidas.....	110
Tabela 2. Índice de Viés Racial Antinegro no Teste Relacional ao Longo dos Sondas e do Follow-up.....	235
Tabela 2. Moda das escolhas e não escolhas no Teste da Boneca para participantes com generalização parcial.....	280
Tabela 3. Análise descritiva do desempenho no FAST ao longo dos diferentes dias de aplicação.....	111
Tabela 3. Percentual de escolhas de faces negras, faces brancas e estímulos abstratos em tentativas com estímulos modelo positivos e negativos.....	231
Tabela 3. Síntese integrativa dos efeitos da intervenção e da convergência entre medidas por participante.....	299
Tabela 4. Correlação entre o Efeito FAST e a Escala Likert em cada momento de aplicação da medida.....	112
Tabela 4. Índice de Viés Racial no Teste da Boneca ao Longo das Sondas e do Follow-up.....	234
Tabela 4. Percentual de escolhas de bonecas negras e brancas na presença de adjetivos positivos e negativos.....	236
Tabela 5. Valência atribuída às faces brancas e negras durante às sondas e follow-up.....	242

Passarelli, D. A. (2026). *Análise do comportamento e preconceito racial: uma história de pesquisa*. (Tese de doutorado, Universidade Federal de São Carlos)

### **Resumo**

Esta tese investiga o viés racial antinegro como responder relacional organizado por contingências sociais, culturais e históricas, propondo uma agenda integrada de análise conceitual, avaliação de medidas explícitas e implícitas e estudos experimentais de intervenção. O Capítulo 1 apresenta o enquadramento teórico da tese, definindo o preconceito racial como responder relacional e discutindo desafios conceituais e metodológicos para sua mensuração e modificação. O Capítulo 2 revisa criticamente a produção empírica sobre viés racial antinegro na Análise do Comportamento, identificando lacunas que orientam os estudos experimentais subsequentes. O Capítulo 3 reúne estudos voltados à avaliação de medidas de viés racial. O Estudo 2, com crianças de 8 a 10 anos, investigou a estabilidade temporal de uma medida explícita de valência (Escala Likert) e do *Function Acquisition Speed Test* (FAST), uma medida implícita, em delineamento de medidas repetidas, evidenciando elevada estabilidade do FAST e sensibilidade à detecção do viés racial antinegro. O Estudo 3, também com crianças de 8 a 10 anos, realizou uma análise multimodal do preconceito racial, identificando efeito substancial de gênero, com desvalorização da mulher negra, e variação contextual da preferência social, indicando controle contextual do responder relacional. O Capítulo 4 concentra os estudos interventivos com crianças, baseados no paradigma da equivalência de estímulos. O Estudo 4, com crianças de 8 a 10 anos, replicou procedimentos descritos por Mizael et al. (2016, 2022), demonstrando formação de classes de equivalência entre faces negras e estímulos positivos, com redução de relações negativas e diminuição das diferenças de valência, porém com generalização limitada e efeitos de deseabilidade social. O Estudo 5, com

crianças de 6 anos, indicou maior convergência entre as medidas e evidências parciais de generalização, embora sem manutenção estável para todos os participantes. No Estudo 6, utilizando procedimentos de equivalência com palavras e acompanhamento longitudinal, gerou efeitos mais robustos de generalização e evidências de validade social, ainda que de forma gradual e heterogênea, com padrões variados de flexibilidade relacional e resistência à mudança. Um achado comum aos Experimentos 3, 4 e 6 foi que a ausência de formação de classes de equivalência esteve associada à ausência de mudanças nas medidas de viés racial. O Capítulo 5 apresenta o estudo com adultos. O Estudo 7, com 116 adultos irlandeses brancos, avaliou os efeitos de diferentes configurações de treino baseadas no FAST sobre medidas implícitas e explícitas, incluindo o jogo do atirador (baseado na Tarefa do Dilema do Policial), a Escala de Racismo Moderno e escalas de valência facial. A condição favorável ao estereótipo aprofundou o viés em relação ao controle, enquanto as condições contra estereotípica e de flexibilidade produziram efeitos mitigadores, sendo a primeira a mais eficaz. Em conjunto, a tese contribui para uma agenda sistemática de pesquisa sobre viés racial antinegro, demonstrando que procedimentos comportamentais podem reorganizar classes previamente estabelecidas, ao mesmo tempo em que evidencia limites e possibilidades da modificação do controle relacional racial ao longo do desenvolvimento e na vida adulta.

*Palavras-chave:* preconceito racial, viés racial antinegro, medidas comportamentais, equivalência de estímulos, flexibilidade relacional, resistência à mudança.

Passarelli, D. A. (2026). *Behavior analysis and racial prejudice: a research history* (Doctoral dissertation, Federal University of São Carlos).

### **Abstract**

This thesis investigates anti-Black racial bias as symbolic responding organized by social, cultural, and historical contingencies, proposing an integrated research agenda that combines conceptual analysis, evaluation of explicit and implicit measures, and experimental intervention studies. Chapter 1 presents the theoretical framework, defining racial prejudice as relational responding and discussing conceptual and methodological challenges in its measurement and modification. Chapter 2 provides a critical review of empirical research on anti-Black racial bias within Behavior Analysis, identifying gaps that guide the subsequent experimental studies. Chapter 3 focuses on the evaluation of racial bias measures. Study 2, conducted with children aged 8 to 10 years, examined the temporal stability of an explicit valence measure (Likert scale) and the *Function Acquisition Speed Test* (FAST), an implicit measure, using a repeated-measures design, demonstrating high stability and sensitivity of the FAST to anti-Black racial bias. Study 3, also with children aged 8 to 10 years, employed a multimodal analysis of racial prejudice, identifying a substantial gender effect characterized by the devaluation of Black women, as well as contextual variation in social preference, indicating contextual control of relational responding. Chapter 4 presents intervention studies with children based on the stimulus equivalence paradigm. Study 4, with children aged 8 to 10 years, replicated procedures described by Mizael et al. (2016, 2022), demonstrating the formation of equivalence classes between Black faces and positive stimuli, with reductions in negative relations and valence differences, albeit with limited generalization and social desirability effects. Study 5, with 6-year-old children, showed greater convergence across measures and partial

generalization, though without stable maintenance for all participants. In Study 6, using word-based equivalence procedures and longitudinal follow-up, produced more robust generalization effects and evidence of social validity, although changes occurred gradually and heterogeneously, with varying patterns of relational flexibility and resistance to change. A common finding across Experiments 3, 4, and 6 was that the absence of equivalence class formation was associated with the absence of change in racial bias measures. Chapter 5 presents adult study. Study 7, conducted with 116 White Irish adults, evaluated the effects of different FAST-based training configurations on implicit and explicit measures, including a shooter task based on the Police Officer's Dilemma Task, the Modern Racism Scale, and facial valence ratings. The stereotype-consistent condition exacerbated bias relative to the control condition, whereas the counter-stereotypical and flexibility conditions produced mitigating effects, with the former being the most effective. Overall, this dissertation contributes to a systematic research agenda on anti-Black racial bias, demonstrating that behavioral procedures can reorganize previously established classes while also delineating the possibilities and limits of modifying racial relational control across development and adulthood.

*Keywords:* racial prejudice, anti-Black racial bias, behavioral measures, stimulus equivalence, relational flexibility, resistance to change.

## **Prefácio**

Esta tese teve origem em uma inquietação científica e ética: como a Análise do Comportamento pode contribuir de forma rigorosa, cumulativa e conceitualmente precisa para a compreensão do preconceito racial, sem reduzir um fenômeno social complexo a explicações simplificadoras ou a intervenções normativas descoladas do controle experimental.

Historicamente, a Análise do Comportamento se desenvolveu, em grande parte, a partir do estudo de processos básicos em organismos não humanos, como ratos e pombos, com resultados frequentemente utilizados para interpretar o comportamento humano. Partia-se da premissa de que os mesmos princípios comportamentais fundamentais operariam em diferentes espécies, o que tornava metodologicamente mais viável investigar tais processos em contextos laboratoriais altamente controlados. Nesse cenário, estudos experimentais com humanos permaneceram relativamente escassos durante as primeiras décadas do desenvolvimento do campo.

A partir das décadas de 1970 e 1980, observa-se um aumento progressivo de pesquisas experimentais com participantes humanos, movimento que levou ao contato com fenômenos que pareciam resistir às explicações disponíveis até então. A equivalência de estímulos emergiu nesse contexto como um fenômeno particularmente relevante, não apenas por sua robustez empírica, mas por sua aparente especificidade ao comportamento humano simbólico. Embora tenham sido propostas tentativas de reduzir a equivalência a processos operantes já conhecidos, essas explicações não se mostraram amplamente convincentes, abrindo espaço para novas formulações conceituais.

A equivalência de estímulos consolidou-se como uma ferramenta central para o estudo de processos simbólicos em humanos, oferecendo uma via experimental para investigar significado, categorização e intercambialidade simbólica. Ao longo das décadas de 1980 e 1990, uma linha consistente de pesquisas demonstrou que a formação de classes de equivalência envolve transferência de função, de modo que estímulos arbitrariamente relacionados passam a adquirir funções semelhantes, inclusive avaliativas e afetivas (Dougher et al., 1994; de Rose et al., 1988; Gandarela et al., 2020).

Esse processo, contudo, não ocorre de maneira indiferenciada. A formação e a reorganização de classes mostraram-se sensíveis à história pré-experimental dos estímulos envolvidos. Estudos como os de Moxon et al. (1993) e Watt et al. (1991) evidenciaram que estímulos afetivamente carregados ou socialmente relevantes tendem a produzir resistência à formação ou reorganização de classes, revelando limites funcionais importantes do procedimento.

A partir desse conjunto de evidências, a equivalência de estímulos passou a oferecer não apenas um modelo experimental do comportamento simbólico, mas também uma base para analisar como significados avaliativos e emocionais são adquiridos, mantidos e transformados por meio de relações arbitrárias. Esse deslocamento abriu caminho para investigações posteriores sobre atitudes, estereótipos e preconceito em contextos sociais. Posteriormente, o desenvolvimento da Teoria das Molduras Relacionais ampliou ainda mais esse horizonte, oferecendo ferramentas conceituais e metodológicas para analisar relações simbólicas complexas para além da equivalência (coordenação), incluindo relações de hierarquia, comparação, oposição e tomada de perspectiva.

Contudo, no que diz respeito especificamente ao racismo antinegro, as contribuições analítico-comportamentais ainda são relativamente poucas e fragmentadas, com um

aumento substancial na última década, especialmente a partir do uso de procedimentos derivados da equivalência de estímulos e da RFT. Esse corpo de pesquisa não pretende substituir ou refazer o que já foi produzido pela psicologia social tradicional, cuja contribuição é extensa e valiosa, mas oferecer uma perspectiva complementar, centrada na análise funcional dos processos comportamentais envolvidos na aquisição, manutenção e transformação do viés racial antinegro.

É nesse contexto histórico e científico que esta tese se insere: como parte de um movimento mais amplo da Análise do Comportamento em direção ao estudo de fenômenos sociais, explorando as oportunidades abertas pela equivalência de estímulos e pela RFT para investigar, de forma experimental e conceitualmente rigorosa, processos envolvidos no preconceito racial.

Ao longo do trabalho, o preconceito racial é tratado não como uma entidade psicológica unitária, nem como avaliação individual, mas como um conjunto de repertórios relacionais historicamente construídos, selecionados e mantidos por contingências culturais. Essa perspectiva desloca o foco da correção de atitudes para a análise dos contextos que organizam o responder, permitindo investigar tanto a aquisição quanto a resistência à mudança de relações simbólicas previamente adquiridas em domínios socialmente relevantes.

Um aspecto central deste trabalho é a recusa de explicações apriorísticas. Em consonância com a postura defendida por Sidman, as interpretações aqui construídas emergem do contato sistemático com os dados, e não da tentativa de acomodar os dados a modelos teóricos previamente estabelecidos. Ao longo da tese, resultados inesperados exigiram reformulações, refinamentos conceituais e, por vezes, a revisão de pressupostos

iniciais. Esse movimento não é apresentado como limitação, mas como parte constitutiva de um empreendimento científico comprometido com a evidência empírica.

A tese se organiza como um programa de investigação progressivo, no qual diferentes estudos exploram, sob condições controladas, como relações simbólicas são estabelecidas, mantidas, flexibilizadas ou resistem à mudança. O uso de múltiplas medidas, delineamentos longitudinais e manipulações contextuais permitiu examinar o preconceito racial como um fenômeno multifacetado, sensível ao contexto e profundamente enraizado em práticas culturais amplas.

O título desta tese dialoga explicitamente com a obra *Equivalence Relations and Behavior: A Research Story* (Sidman, 1994), indicando a opção por uma forma de apresentação que concebe a produção científica como um percurso investigativo, e não apenas como a exposição linear de resultados. Em consonância com essa inspiração, a estrutura do trabalho foi organizada de modo a tornar explícitas as decisões metodológicas, as dúvidas analíticas, os ajustes procedimentais e as implicações conceituais que acompanharam o desenvolvimento da pesquisa.

Tal escolha estrutural reflete uma afinidade com a perspectiva de Sidman, segundo a qual o conhecimento científico se constrói progressivamente a partir do contato sistemático com os dados, permitindo que observações empíricas orientem, de forma cumulativa, o refinamento conceitual. Trata-se, portanto, de um diálogo metodológico e epistemológico, sem a pretensão de equiparação quanto ao escopo, ao contexto histórico ou à contribuição da obra que lhe serve de referência.

Nesse sentido, cada capítulo empírico é antecedido por um *background* e seguido por um comentário analítico. A função desses elementos não é repetir o conteúdo dos artigos, mas situá-los conceitual e historicamente dentro do programa de pesquisa mais

amplo desenvolvido nesta tese. Em artigos científicos, grande parte do percurso experimental e conceitual se perde em função das exigências de uma estrutura protocolar rígida, que privilegia concisão, padronização e aderência a formatos editoriais. Decisões metodológicas, impasses interpretativos, hipóteses abandonadas e implicações teóricas mais amplas frequentemente não encontram espaço nesse formato, apesar de serem centrais para a compreensão do fenômeno investigado.

Os *backgrounds* cumprem, portanto, a função de contextualizar cada estudo, explicitando as perguntas que o motivaram, as limitações dos estudos anteriores e as razões para as escolhas metodológicas adotadas. Já os comentários têm como objetivo integrar os achados empíricos a uma análise conceitual mais ampla, discutindo seus limites, suas contribuições e suas implicações para o entendimento do preconceito racial a partir da análise do comportamento. Essa estrutura busca preservar o rigor científico dos artigos, sem renunciar a uma reflexão mais profunda sobre os processos investigados.

Cabe destacar que a tese está organizada em cinco capítulos, de modo que o **Capítulo 1** apresenta uma introdução teórica ao fenômeno de interesse, situando o preconceito racial antinegro no âmbito da análise do comportamento, com ênfase em processos simbólicos, controle de estímulos e responder relacional. O **Capítulo 2** oferece um panorama geral das pesquisas empíricas sobre viés racial antinegro, discutindo medidas, procedimentos e tendências metodológicas na literatura. O **Capítulo 3** reúne estudos empíricos voltados à investigação de medidas de viés racial em crianças, examinando estabilidade, convergência entre medidas e efeitos de desejabilidade social. O **Capítulo 4** apresenta intervenções experimentais com crianças, explorando procedimentos baseados em equivalência de estímulos e seus efeitos sobre diferentes dimensões do responder relacional racialmente enviesado. Por fim, o **Capítulo 5** descreve uma intervenção com

adultos, examinando flexibilidade relacional, resistência à mudança e efeitos de diferentes arranjos de contingências.

Cabe destacar que o programa de pesquisa desenvolvido nesta tese seguiu uma lógica cumulativa, na qual cada estudo foi orientado, ao menos em parte, pelos resultados do anterior. O Estudo 1 mapeou o estado da produção empírica na área e identificou lacunas que justificaram os estudos subsequentes. Os Estudos 2 e 3 responderam a uma questão metodológica prévia à intervenção: as medidas disponíveis são estáveis e sensíveis o suficiente para detectar o fenômeno e avaliar mudanças? Somente após estabelecer essa base é que os estudos interventivos foram iniciados. O Estudo 4 replicou o procedimento de Mizael et al. (2016) como passo necessário antes de qualquer comparação ou extensão. Os Estudos 5 e 6 responderam a perguntas abertas pelo Estudo 4: os efeitos generalizam para crianças mais jovens? O Estudo 7, por fim, deslocou o foco para adultos, motivado pelos dados de resistência à mudança observados nos estudos com crianças. Esse percurso não foi inteiramente planejado a priori: em vários momentos, resultados inesperados reorientaram as decisões seguintes, o que é explicitado nos comentários analíticos ao longo dos capítulos.

Por fim, este trabalho não pretende oferecer soluções universais nem receitas de intervenção. Seu objetivo é mais modesto e, ao mesmo tempo, mais exigente: identificar processos, delimitar condições, explicitar limites e ampliar o repertório analítico disponível para o estudo do preconceito racial a partir da Análise do Comportamento. Se esta tese contribuir para a construção de agendas de pesquisa mais integradas, conceitualmente precisas e empiricamente fundamentadas, terá cumprido seu propósito

## Capítulo 1. Introdução

### **Comportamento Simbólico, Atitudes e Preconceito Racial: Uma Caixa de Pandora para a Análise do Comportamento?**

A ciência do comportamento construiu, ao longo de sua história, modelos robustos para explicar como organismos aprendem a agir no mundo, desde respostas reflexas organizadas por relações respondentes (Pavlov, 1927; Watson, 1913) até o comportamento operante, selecionado e mantido por suas consequências (Skinner, 1953, 1957). No entanto, quando aplicados a fenômenos socialmente sensíveis, como atitudes e preconceito racial, esses mesmos modelos passam a revelar seus limites e tensões.

Essas tensões emergem, em grande medida, tanto da dificuldade de analisar funcionalmente o fenômeno em questão quanto do caráter mentalista historicamente associado à própria nomenclatura utilizada para descrevê-lo. Diferentemente de repertórios organizados por relações funcionais observáveis e identificáveis, o preconceito racial se estrutura em redes simbólicas amplas, historicamente acumuladas e culturalmente mantidas, nas quais avaliações, categorias sociais, expectativas e padrões de ação se articulam de forma densa (Ryle & Tanney, 2009). Ao tentar apreender esse domínio, a análise do comportamento se depara com um campo no qual múltiplos controles de estímulo, relações de coordenação, comparação e hierarquia, regras culturais e práticas sociais operam simultaneamente, dificultando a delimitação de unidades funcionais (Passarelli et al., 2023).

É nesse sentido que o preconceito racial pode ser compreendido como uma verdadeira “Caixa de Pandora” para a análise do comportamento. Não porque sua investigação conduza a efeitos indesejáveis, mas porque, ao se examinar suas bases comportamentais, impõe-se o desafio analítico de definir funcionalmente fenômenos que se

organizam em múltiplos níveis. Nesse domínio, processos básicos de aprendizagem interagem com histórias verbais extensas, relações simbólicas altamente gerativas e contingências culturais amplas, de modo que qualquer proposta de intervenção requer, como condição inicial, a identificação das funções comportamentais específicas envolvidas, e não apenas de suas topografias ou descrições avaliativas.

Apesar das tensões e dificuldades pontuadas, investigar esse fenômeno complexo e multifacetado não exige o abandono da análise do comportamento. No entanto, faz-se necessária a integração de diferentes abordagens desenvolvidas no interior da própria área. Princípios básicos de aprendizagem, análise do comportamento verbal, equivalência de estímulos, responder relacional e práticas culturais precisam ser articulados para compreender como repertórios individuais operam em contextos sociais amplamente estruturados para manter desigualdades.

Nesse sentido, a metáfora da Caixa de Pandora não remete apenas à complexidade do fenômeno, mas também à possibilidade de intervenção. Assim como no mito, em que a esperança permanece após a liberação dos males, a análise do comportamento oferece ferramentas conceituais e metodológicas para identificar pontos de mudança possíveis, ainda que parciais, em sistemas simbólicos e culturais altamente estáveis.

### **Racismo e Viés Racial**

A compreensão do *preconceito racial* exige distinguir, de forma analiticamente rigorosa, entre repertórios individuais de resposta e práticas sociais coordenadas sustentadas por contingências culturais e relações assimétricas de poder. Embora atitudes, estereótipos e vieses possam ser observados no comportamento de indivíduos, o racismo não se reduz à soma desses repertórios individuais. Ele envolve sistemas historicamente

organizados de práticas sociais, controle institucional e distribuição desigual de recursos entre grupos sociais (Mizael & Sampaio, 2017; Saini & Vance, 2020).

O racismo pode ser compreendido como um conjunto de relações sistemáticas de poder, historicamente organizadas, por meio das quais grupos hegemônicos mantêm controle assimétrico sobre recursos materiais, simbólicos e sociais em detrimento de grupos racialmente marginalizados (Almeida, 2019; Saini & Vance, 2020; Mizael & Sampaio, 2019). Essas relações não se sustentam apenas por ações individuais isoladas ou por hostilidade explícita, mas por práticas sociais coordenadas que produzem benefícios consistentes para grupos dominantes, legitimando e naturalizando hierarquias raciais ao longo do tempo (Munanga, 2004; Passarelli et al., 2023). Assim, o racismo opera como um fenômeno cultural, mantido por contingências sociais amplas, controle institucional e relações assimétricas de poder, que estabilizam arranjos sociais desiguais.

Central para a compreensão do racismo está a própria construção do conceito de raça. Embora historicamente apresentado como uma categoria natural, a raça não corresponde a nenhuma divisão biologicamente significativa da espécie humana. Evidências acumuladas nas ciências naturais demonstram que características fenotípicas, como cor da pele ou textura do cabelo, variam de forma contínua e não delimitam categorias biológicas discretas (Lewontin, 1972; Jablonski, 2012; Smedley & Smedley, 2005). A ausência de fronteiras biológicas estáveis indica que a raça não pode ser compreendida como um “tipo natural”, mas como um produto de processos históricos, sociais e políticos de classificação, articulados para organizar e justificar desigualdades (Du Bois, 1903; Memmi, 2000; Munanga, 2004; Almeida, 2019; Passarelli et al., 2023).

O caráter não natural dessas categorias torna-se especialmente evidente em sua instabilidade contextual. Se a raça refletisse distinções biológicas, as classificações raciais

permaneceriam consistentes ao longo do tempo e entre contextos culturais. No entanto, as mesmas características fenotípicas podem ser interpretadas e classificadas de maneiras distintas em diferentes países, culturas e situações sociais. Um indivíduo classificado como branco no contexto brasileiro pode ser categorizado como “não europeu” ou “latino” em outro, sem qualquer alteração em suas características biológicas. Essa variabilidade evidencia que a variável operante não é a biologia, mas o arcabouço simbólico de classificação compartilhado pela comunidade verbal (Haney López, 1995; Smedley & Smedley, 2005; Passarelli et al., 2023).

Do ponto de vista da análise do comportamento, essas classificações são mantidas por práticas culturais, entendidas como padrões recorrentes de comportamento verbal e não verbal selecionados e mantidos por suas consequências sociais agregadas (Glenn, 1988, 1991). Categorias raciais funcionam, assim, como estímulos verbais convencionais, cujas propriedades comportamentais são adquiridas por meio de histórias de reforçamento, regras culturalmente transmitidas e controle institucional (Parks & Kirby, 2022; Seavon, 2022). Por meio dessas práticas, características fenotípicas passam a funcionar como pistas para respostas diferenciais, organizando redes simbólicas que regulam julgamentos, emoções e ações mesmo na ausência de contingências diretas ou de experiências interpessoais explícitas (Mizael & de Rose, 2016; Passarelli et al., 2023).

Pesquisas sobre equivalência de estímulos demonstram que, uma vez estabelecidas relações entre estímulos arbitrários em uma comunidade verbal, novas relações podem ser derivadas sem reforçamento direto, permitindo que funções avaliativas e emocionais sejam transferidas entre elementos da rede simbólica (Sidman, 1994; de Rose & Bortolotti, 2007). Nesse processo, rótulos raciais, características físicas e descritores avaliativos podem tornar-se funcionalmente equivalentes, de modo que funções atribuídas a um elemento se

estendam aos demais. A transformação de funções permite que padrões de resposta racializados sejam mantidos mesmo na ausência de contato direto com membros dos grupos alvo, contribuindo para a estabilização de práticas sociais no nível coletivo (Passarelli et al., 2023).

A Teoria das Molduras Relacionais amplia essa análise ao descrever como diferentes tipos de relações simbólicas, como hierarquia, oposição, comparação e distinção são aprendidos por meio de contingências verbais socialmente mediadas (Hayes et al., 2001; Barnes-Holmes et al., 2017). Por meio dessas molduras, estímulos raciais passam a participar de redes relacionais altamente gerativas, organizando distinções do tipo “nós–eles” e hierarquias sociais que ajudam a explicar a persistência do racismo mesmo em contextos que explicitamente endossam valores igualitários (Barnes-Holmes et al., 2010; Devine et al., 2012; FitzGerald & Hurst, 2017).

Nesse panorama, torna-se fundamental distinguir o racismo do viés racial. Enquanto o racismo refere-se a práticas sociais coordenadas, sustentadas por contingências culturais amplas, controle institucional e relações assimétricas de poder, o viés racial diz respeito a padrões diferenciais de responder observáveis no comportamento individual, expressos em avaliações, escolhas e tomadas de decisão (Passarelli et al., 2023). Do ponto de vista comportamental, o viés racial pode ser compreendido como um repertório adquirido sob controle de redes simbólicas culturalmente estabelecidas, nas quais grupos sociais são organizados relacionalmente por coordenação, comparação, oposição etc. Como resultado, o responder passa a apresentar direcionalidade, isto é, uma inclinação sistemática que pode ser tanto uma favorabilidade dirigida ao grupo dominante, quanto uma desfavorabilidade dirigida a grupos marginalizados, organizando respostas do tipo *pró–contra*.

Essa distinção entre níveis de análise do individual e grupo é central para compreender os limites e alcances da pesquisa experimental. O viés racial, por operar no nível individual, é frequentemente acessível à investigação empírica e à modificação por meio de procedimentos experimentais. O racismo, por sua vez, opera como um sistema de práticas culturais que tende a neutralizar ou limitar os efeitos de mudanças individuais, especialmente quando as contingências institucionais e culturais que o sustentam permanecem inalteradas (Lai et al., 2014, 2016; Sawyer & Gampa, 2018; Passarelli et al., 2023).

É nesse ponto que o racismo se torna uma das dimensões mais críticas da metáfora da Caixa de Pandora adotada nesta tese. Mudanças no repertório individual passam a operar em um contexto amplamente estruturado para manter desigualdades raciais, no qual práticas sociais, regras culturais e controles institucionais tendem a neutralizar, limitar ou reverter seus efeitos. A interação entre mudança individual e contexto social passa, assim, a constituir uma variável central para a análise do comportamento em fenômenos socialmente complexos.

### **Perspectivas da Análise do Comportamento sobre Atitudes, Preconceito, Preferências e Discriminação**

A psicologia social experimental ofereceu contribuições relevantes para a compreensão das atitudes, estereótipos e do preconceito. No entanto, essa abordagem frequentemente recorre a explicações mentalistas, atribuindo status causal a eventos internos hipotéticos, como atitudes, crenças ou disposições psicológicas. Nessa perspectiva, o preconceito tende a ser concebido como algo “interno” ao indivíduo, ao qual se atribui a função explicativa de comportamentos racistas observáveis (Guerin, 2005). Como consequência, muitas intervenções direcionam-se à modificação desses supostos estados

internos, em detrimento da análise das condições ambientais, históricas e culturais que organizam e mantêm tais padrões comportamentais.

Tradicionalmente, o preconceito racial tem sido definido como atitudes, isto é, avaliações negativas dirigidas a um grupo racial ou a seus membros (Mizael & de Rose, 2017; Nogueira, 2006; Passarelli, Mizael et al., 2024). Os estereótipos, por sua vez, referem-se à atribuição de características generalizadas e simplificadas a grupos sociais, inclusive àqueles com os quais não houve contato direto (Stangor, 2009).

Do ponto de vista comportamental, essas duas classes de fenômenos podem ser distinguidas funcionalmente. Os estereótipos podem ser compreendidos como produtos verbais de uma comunidade verbal, nos quais categorias sociais passam a ser descritas de forma recorrente por determinados atributos, independentemente de experiências diretas com os indivíduos descritos (Parks & Kirby, 2022; Seavon, 2022). Nessa perspectiva, funcionam como regras, isto é, estímulos verbais que especificam relações entre categorias e atributos e controlam o comportamento de quem os segue. Esses repertórios são mantidos por práticas verbais culturalmente compartilhadas e por consequências sociais generalizadas. O preconceito, por sua vez, refere-se ao padrão avaliativo que se organiza a partir desses repertórios descritivos. Nesses casos, estímulos que sinalizam pertencimento categorial passam a controlar respostas de viés racial, frequentemente caracterizadas por desfavorabilidade dirigida a grupos minoritários. Um exemplo é o estereótipo que associa pessoas negras à criminalidade, o qual tende a organizar avaliações negativas como “agressivo”, “perigoso” ou “mal-intencionado”, mesmo na ausência de experiências diretas.

Essas definições são úteis na medida em que enfatizam o componente avaliativo do preconceito e permitem sua mensuração sistemática, além de possibilitarem diálogo com explicações comportamentais. No entanto, do ponto de vista funcional, permanecem

parciais: embora sejam eficientes para descrever a topografia das respostas avaliativas, deixam de lado dimensões centrais para a compreensão do fenômeno, como o controle contextual, as relações de preferência entre categorias, os controles de estímulos envolvidos no responder e, sobretudo, as funções comportamentais que essas avaliações passam a exercer. Ao restringir o preconceito à dimensão avaliativa atribuída a grupos sociais, tais definições pouco esclarecem como essas avaliações são aprendidas, organizadas e mantidas em redes simbólicas mais amplas. É nesse ponto que o estudo do responder relacional oferece uma ampliação conceitual necessária.

Contudo, antes de abordar o responder relacional propriamente dito, é importante ressaltar que atitudes e preconceitos são comportamentos aprendidos. Processos básicos de aprendizagem, como condicionamento respondente, condicionamento operante, generalização, controle de estímulos, comportamento verbal e aprendizagem observacional, podem estabelecer diretamente respostas avaliativas desfavoráveis dirigidas a grupos sociais (Belisle et al., 2022; Catania, 2017; Matsuda et al., 2020). No entanto, pesquisas sobre equivalência de estímulos e, posteriormente, sobre a Teoria das Molduras Relacionais (RFT; Hayes et al., 2001) ampliaram substancialmente a compreensão desses fenômenos ao demonstrar como novas relações podem emergir sem treino direto.

Como mencionado na seção anterior, estudos sobre equivalência de estímulos (Sidman, 1971, 1994) mostram que, a partir do treino de um conjunto pequeno de relações condicionais, relações não diretamente ensinadas podem ser derivadas, de modo que estímulos distintos passam a integrar uma mesma classe funcional, tornando-se intercambiáveis entre si e permitindo a transferência de funções avaliativas, emocionais ou motivacionais (Barnes-Holmes et al., 2000; Bortoloti & de Rose, 2009; Silveira et al., 2018; de Rose, 2022). Esse conjunto de achados fornece um modelo comportamental

parcimonioso para compreender o viés racial, as atitudes, os estereótipos e os preconceitos, ao mostrar como avaliações podem ser mantidas e generalizadas na ausência de contato direto, com base apenas na pertença categorial (Grey & Barnes, 1996). Nessa perspectiva, o preconceito racial corresponde a um padrão de responder relacional enviesado, no qual atributos étnicos passam a controlar avaliações negativas e respostas desfavoráveis, conferindo ao viés racial uma direcionalidade desfavorável.

Além disso, evidências empíricas destacam o papel do controle contextual na organização dessas respostas. O pertencimento a uma categoria social pode funcionar como contexto discriminativo, alterando o significado funcional de estímulos ou comportamentos topograficamente idênticos, de modo que a avaliação produzida depende de quem os emite (Kohlenberg et al., 1991). Por exemplo, um mesmo comportamento, como correr em via pública, tende a ser interpretado como exercício físico quando emitido por um indivíduo branco e como fuga ou ameaça quando emitido por um indivíduo negro. Outro aspecto relevante do controle contextual diz respeito à organização de respostas com topografia aparentemente positiva dirigidas a determinados grupos, mas cujas funções contribuem para a manutenção de desigualdades e privilégios (Passarelli et al., 2023; Passarelli, Roche et al., 2024). Nesses casos, atributos socialmente valorizados podem operar como estereótipos positivos apenas em nível descritivo, enquanto produzem consequências restritivas no plano funcional. Por exemplo, a associação de pessoas negras a características como “força física” ou “habilidade na cozinha” pode parecer favorável em termos avaliativos, mas frequentemente sustenta expectativas limitadas de desempenho e oportunidades, como a crença de que indivíduos negros seriam mais adequados apenas a trabalhos físicos ou de baixa qualificação.

A partir desses achados, é possível identificar ao menos duas características fundamentais dos preconceitos sob a ótica do responder relacional: (1) a equivalência entre categorias sociais e atributos avaliativos, frequentemente negativos, com generalização desses atributos a todos os membros da categoria; e (2) o papel da categoria como contexto, alterando o significado funcional de estímulos e ações que, em outros contextos, não evocariam avaliações negativas. Esses processos não se restringem a categorias raciais e podem ser observados em diferentes domínios sociais, incluindo o ambiente acadêmico, onde determinadas abordagens teóricas ou grupos profissionais também podem tornar-se alvos de avaliações desfavoráveis generalizadas.

A Teoria das Molduras Relacionais (*Relational Frame Theory* – RFT) amplia esse modelo ao descrever outros tipos de relações além da equivalência (que na RFT é muitas vezes chamada de coordenação), como oposição, comparação, hierarquia e dêiticas. Essas molduras permitem compreender como o viés envolve não apenas avaliações desfavoráveis dirigidas a um grupo, mas também favorabilidade concomitante dirigida a outro, organizando respostas do tipo *pró–contra*. Ao estruturar relações comparativas e hierárquicas entre categorias sociais, essas molduras introduzem a dimensão das *preferências raciais*, que, quando estabilizadas socialmente, tendem a traduzir-se em distribuições desiguais de recursos simbólicos, sociais e materiais. Além disso, molduras dêiticas, relacionadas às perspectivas de “eu–outro”, “aqui–ali” e “agora–então”, têm sido associadas à tomada de perspectiva e à empatia. Evidências sugerem que respostas empáticas são mais prováveis quando o outro é percebido como semelhante, isto é, quando há coordenação relacional entre valores, crenças e características compartilhadas (Suarez et al., 2024)

Portanto, avançando nessa discussão, é possível identificar outras três dimensões centrais do preconceito racial sob a ótica do comportamento simbólico: (1) padrões de favorabilidade dirigidos a grupos dominantes, que organizam preferências e controlam a distribuição desigual de recursos simbólicos, sociais e materiais; (2) relações simbólicas de comparação, oposição e hierarquia entre grupos raciais, que estruturam distinções do tipo “nós–eles” e sustentam desigualdades ao estabelecer valor diferencial entre categorias sociais; e (3) assimetrias na tomada de perspectiva social, nas quais respostas empáticas tendem a ocorrer com maior probabilidade em relação a membros do próprio grupo, especialmente quando há coordenação de valores, crenças e características compartilhadas. Esses processos operam de forma articulada e não se restringem a contextos explicitamente raciais, podendo ser observados em diversos domínios sociais nos quais categorias grupais organizam avaliações, preferências e *práticas discriminatórias*.

Nesse sentido, pode-se identificar mais uma dimensão do preconceito racial: a discriminação racial, compreendida como um padrão de responder relacional diferencial sob controle de atributos raciais, no qual a pertença categorial aumenta a probabilidade de consequências desfavoráveis dirigidas a membros de grupos raciais. Diferentemente das preferências, que operam predominantemente por reforço diferencial positivo, expresso na alocação favorecida de recursos, oportunidades ou atenção a membros de grupos dominantes, a discriminação envolve tratamento diferencial com consequências aversivas ou processos de extinção, tais como exclusão, restrição, vigilância excessiva, desconfiança ou punição (Catania, 2014; Mizael & Sampaio, 2019). Em ambos os casos, a pertença racial funciona como condição sob a qual padrões específicos de resposta são selecionados e mantidos. Esses repertórios não dependem de contato direto, mas de práticas sociais recorrentes e histórias verbais culturalmente estabelecidas, de modo que comportamentos

topograficamente semelhantes produzem consequências sistematicamente distintas conforme o grupo racial ao qual o indivíduo é atribuído. Assim, preferências e discriminação configuram formas complementares de tratamento diferencial, que, ao operar conjuntamente, organizam a distribuição desigual de recursos, oportunidades e sanções no cotidiano social.

### **A Caixa de Pandora do Preconceito Racial: Males Persistem, a Esperança Permanece**

Diante desse panorama, esta tese parte do pressuposto de que compreender o preconceito racial a partir da análise do comportamento exige não apenas delimitar conceitualmente seus componentes, mas também examinar como esses componentes têm sido operacionalizados, medidos e modificados empiricamente ao longo do desenvolvimento da área. Ao longo dos capítulos seguintes, sustenta-se que o avanço conceitual da análise do comportamento no estudo do preconceito racial depende, de forma indissociável, do desenvolvimento de uma agenda empírica consistente, na qual refinamentos teóricos caminham conjuntamente com a evolução dos procedimentos de mensuração e das estratégias de intervenção, especialmente aquelas fundamentadas no comportamento simbólico e no responder relacional.

Assim, esta tese organiza-se em torno de três eixos articulados: uma revisão crítica sobre trabalhos em preconceito racial antinegro na análise do comportamento; a análise crítica das medidas comportamentais utilizadas para avaliar viés racial e preconceito; o exame das intervenções experimentais desenvolvidas para alterar esses repertórios; e a discussão de como os avanços empíricos nesses domínios têm tensionado, refinado e ampliado os próprios conceitos comportamentais empregados para descrever o fenômeno. Ao fazê-lo, busca-se mostrar que o estudo do preconceito racial não apenas desafia a análise do comportamento, mas também constitui um campo privilegiado para observar, de

forma concreta, a evolução de seus modelos explicativos frente a fenômenos simbólicos complexos.

Portanto, a seção seguinte inicia esse percurso ao revisar os principais procedimentos empregados para mensurar o preconceito e o viés racial, estabelecendo as bases conceituais e metodológicas que sustentam as investigações empíricas desenvolvidas ao longo desta tese.

####

## **Capítulo 2. Análise do Comportamento, Racismo e Preconceito Racial: Onde estivemos e para onde vamos?**

### **Background**

Em novembro de 2025, estive na Universidade de Edimburgo para uma visita técnica ao laboratório da professora Táhcita Mizael, referência internacional em estudos sobre raça e gênero no contexto da Análise do Comportamento.

Essa visita insere-se em uma trajetória acadêmica construída a partir de um programa de pesquisa desenvolvido no Laboratório de Estudos sobre Comportamento Humano (LECH), sob a orientação do professor Júlio de Rose. As primeiras investigações empíricas sobre preconceito racial nesse contexto foram conduzidas por Marília de Carvalho, no âmbito de seus estudos de doutorado. Posteriormente, essa agenda de pesquisa foi ampliada e aprofundada por Táhcita, em seu mestrado e doutorado, também sob orientação do professor Júlio de Rose. E, em sequência, essa agenda se tornou parte do meu doutorado, sob orientação do professor Júlio de Rose e coorientação de Táhcita e do professor Bryan Roche.

Durante uma reunião online de planejamento da visita, manifestei à Táhcita meu interesse em realizar uma revisão que reunisse todos os estudos empíricos sobre preconceito racial antinegro na Análise do Comportamento. Esse interesse não surgiu de forma abstrata, mas a partir de um incômodo progressivo e persistente. Em particular, chamou-me a atenção a edição especial de 2022 da *Behavior Analysis in Practice* dedicada ao antirracismo. Embora a iniciativa tenha sido, sem dúvida, relevante como intervenção em nível cultural e como estratégia de visibilização do tema, a leitura dos artigos revelou um padrão que me causou inquietação: a maioria dos trabalhos era de natureza teórica, oferecendo análises conceituais e propostas de intervenção, mas com escassa produção

empírica e, frequentemente, sem definições funcionais claras de termos como racismo, preconceito racial ou viés racial.

Observava-se, nesses textos, o uso recorrente e confiante de expressões como *racismo sistêmico*, *racismo institucional* ou *racismo estrutural*, muitas vezes sem que fosse explicitado quais repertórios estavam sendo descritos, quais unidades comportamentais estavam em jogo ou quais variáveis efetivamente poderiam ser mensuradas e manipuladas experimentalmente. Esse cenário levantou uma questão central para mim: como avançar teoricamente em um domínio no qual os próprios fenômenos permanecem pouco delimitados do ponto de vista empírico?

Como resposta inicial a essa inquietação, escrevi um artigo teórico no qual propus definições comportamentais de raça e racismo, bem como uma divisão analítica de unidades funcionais para a Análise Comportamental do Racismo, distinguindo racismo estrutural, racismo sistêmico (institucional) e racismo individual. Esse trabalho foi publicado na *Revista Brasileira de Análise do Comportamento* (REBAC) e contou com cuidadosa edição da professora Camila Muchon de Melo. Apesar de reconhecer a importância desse artigo como parte do meu percurso formativo, não o considero um ponto de chegada. Ele foi escrito em um momento em que minha maturidade teórica sobre o tema ainda era limitada, e quando meus estudos experimentais estavam em fase inicial, marcados por tentativas exploratórias e pilotos que não produziram os resultados esperados.

Ainda assim, esse artigo desempenhou um papel fundamental: tornou explícita, para mim, a necessidade incontornável de dados empíricos para sustentar definições, modelos teóricos e propostas de intervenção. Mais do que isso, evidenciou um risco recorrente no campo: o de inverter a ordem do desenvolvimento científico, formulando teorias amplas antes de compreender, de forma sistemática, os repertórios que se pretende explicar. Essa

constatação me levou a refletir que, em muitos trabalhos sobre racismo na Análise do Comportamento, estávamos trilhando um caminho inverso àquele proposto por Sidman: partir de construções conceituais robustas sem que houvesse, previamente, uma base empírica suficientemente consolidada.

É verdade que a estratégia aqui adotada pode parecer, à primeira vista, paradoxal. Afinal, respondo a essa inquietação justamente com uma revisão sistemática, isto é, com um estudo secundário sem coleta de dados com participantes. No entanto, esse aparente paradoxo se desfaz quando se reconhece que a finalidade da revisão não é substituir a pesquisa experimental, mas mapear rigorosamente o que já foi feito empiricamente, identificar lacunas e delimitar com precisão onde a Análise do Comportamento esteve e, sobretudo, para onde ela pode e precisa avançar. Defendo, ao longo deste trabalho, que uma agenda experimental estruturada é indispensável para o progresso do campo e que nossas pressuposições teóricas devem estar ancoradas em evidências observáveis e replicáveis.

Sem mais delongas, o primeiro artigo que compõe esta tese consiste em uma revisão sistemática de estudos empíricos que abordaram preconceito racial ou racismo antinegro no âmbito da Análise do Comportamento. O recorte foi deliberadamente delimitado para permitir acesso a todos os trabalhos empíricos disponíveis nessa temática, com o objetivo de compreender o estado atual da produção científica da área, identificar padrões metodológicos, conceituais e analíticos, e estabelecer bases mais sólidas para a construção de uma agenda de pesquisa que articule, de forma indissociável, avanço teórico e investigação empírica.

\*\*\*

## **Estudo 1 - Uma Revisão Sistemática de Dados Empíricos sobre Viés Racial Antinegro na Análise do Comportamento<sup>1</sup>**

Apesar da crescente atenção à diversidade, equidade e inclusão (DEI), a pesquisa empírica sobre o viés racial antinegro na análise do comportamento permanece limitada. Esta revisão sistemática, conduzida segundo as diretrizes PRISMA, examinou 20 estudos empíricos publicados em 16 periódicos da área entre 2007 e 2024. Os estudos foram organizados em três domínios metodológicos: (1) levantamentos transversais e correlacionais, (2) estudos de mensuração do viés racial antinegro e (3) intervenções experimentais. Os estudos transversais e correlacionais examinaram relações entre valores, flexibilidade psicológica, saúde mental e atitudes raciais. Os estudos de mensuração utilizaram principalmente instrumentos comportamentais baseados na Teoria das Molduras Relacionais, como o IRAP e o FAST, os quais identificaram vieses implícitos frequentemente não detectados por medidas de autorrelato. As intervenções, especialmente aquelas baseadas em treino de equivalência, apresentaram efeitos promissores, ainda que preliminares, na alteração de redes relacionais enviesadas, tanto em crianças quanto em adultos, com destaque para procedimentos envolvendo clarificação de valores e pareamentos relacionais inconsistentes. Em todas as categorias, foram observadas limitações metodológicas recorrentes, incluindo amostras ocidentalizadas e ricas, ausência de avaliação de viés de não resposta e escassez de testes de seguimento e generalização. Esses achados apontam para a necessidade de uma agenda de pesquisa mais cumulativa, conceitualmente precisa e funcionalmente fundamentada para sustentar esforços antirracistas na análise do comportamento.

---

<sup>1</sup> Tradução do manuscrito de Passarelli, D. A., Mizael, T. M., & De Rose, J. C. (2026). *A systematic review of empirical data on anti-Black racial bias in behavior Analysis*. [Manuscrito submetido a periódico científico, atualmente em processo de revisão].

## Introdução

O preconceito racial, particularmente o preconceito antinegro, permanece como um dos fenômenos mais persistentes e socialmente prejudiciais da sociedade. Embora a psicologia estude o preconceito racial extensivamente desde meados do século XX, a análise do comportamento apenas recentemente começou a se engajar com esse tema crítico. Estudos fundacionais, como os experimentos das bonecas de Clark e Clark (1947), demonstraram como a segregação racial distorcia a autopercepção de crianças negras, estabelecendo as bases para a pesquisa sobre atitudes raciais. Pesquisas no campo do desenvolvimento também demonstraram que as atitudes raciais emergem precocemente e são moldadas por processos de aprendizagem social. Por exemplo, Aboud et al. (2003) constataram que o preconceito racial entre crianças estava associado a um aumento na exclusão de pares pertencentes a outros grupos raciais.

No início dos anos 2000, a psicologia avançou em modelos teóricos de viés implícito, especialmente por meio de paradigmas de condicionamento avaliativo (De Houwer, Thomas & Baeyens, 2001; Hofmann et al., 2010). Esses modelos demonstraram como atitudes, especialmente atitudes raciais, podem ser moldadas por pareamentos repetidos entre estímulos sociais (por exemplo, rostos negros) e estímulos com valência positiva ou negativa, mesmo na ausência de relatos verbais explícitos que descrevam os pareamentos estímulo–estímulo relevantes. Pesquisas em condicionamento avaliativo mostraram como tais pareamentos podem tanto estabelecer quanto modificar avaliações raciais (Olson & Fazio, 2001, 2006; Kurdi & Banaji, 2017; Zanon, De Houwer & Gast, 2014), destacando o

caráter automático, sensível ao contexto e passível de modificação das avaliações raciais.

Apesar desse corpo de conhecimento em expansão, a análise do comportamento permaneceu historicamente desconectada dessa literatura, em razão de sua ênfase no comportamento diretamente observável e de sua rejeição a construtos mentalistas, como “atitudes” (Baer et al., 1968; Skinner, 1953). Embora metodologicamente fundamentada, essa hesitação também refletiu a dificuldade de definir fenômenos complexos como o racismo em termos funcionais e de isolar suas variáveis de controle em contextos experimentais (Passarelli et al., 2023).

Essa lacuna histórica começou a ser reduzida com desenvolvimentos teóricos no interior da própria análise do comportamento. A equivalência de estímulos (Sidman, 1994) e, posteriormente, a Teoria das Molduras Relacionais (*Relational Frame Theory* – RFT; Hayes et al., 2001) ofereceram ferramentas para analisar o preconceito racial como padrões de relações verbais aprendidas e funções de estímulos derivadas (Mizael & de Rose, 2016). Esses referenciais permitiram que analistas do comportamento abordassem o viés racial sem abandonar uma estrutura analítico-funcional.

### **Definições Comportamentais de Viés Racial**

Como mencionado, os avanços teóricos na análise do comportamento ofereceram uma alternativa às explicações mentalistas frequentemente encontradas na psicologia social (Watters et al., 2023). Em vez de atribuir o viés racial a estados internos ou crenças cognitivas, a análise do comportamento o explica em termos de relações funcionais entre estímulos (Passarelli, Roche et al.,

2024; Passarelli, Mizael et al., 2024; Mizael & de Rose, 2016) e do impacto de contingências contextuais (Guerin, 2003) e culturais sobre o comportamento (Saini & Vance, 2020).

Assim, sob uma perspectiva comportamental, o viés racial pode ser definido como um padrão de respostas relacionais no qual indivíduos respondem de forma mais favorável (ou positiva) ou desfavorável (ou negativa) a estímulos relacionados a categorias raciais (ver Mizael & de Rose, 2016; ver também Passarelli et al., 2023; Passarelli, Mizael et al., 2024). Essas respostas são moldadas tanto por histórias de reforçamento direto quanto por respostas relacionais derivadas, frequentemente envolvendo molduras relacionais como coordenação (por exemplo, “negro” = “ruim”), comparação (por exemplo, “branco é melhor do que negro”) e/ou oposição (por exemplo, “se branco é bom, então negro é ruim”).

É importante destacar que estímulos relacionados a grupos raciais, como tom de pele, nomes ou características faciais, podem adquirir funções avaliativas (por exemplo, “bom” ou “ruim”) por meio de processos de aprendizagem socialmente mediados. Essas respostas relacionais derivadas podem influenciar o comportamento de maneiras significativas, orientando ações como aproximação ou esquiva (Lindström et al., 2014), preferências (Dehon et al., 2017; Torrez et al., 2024) ou comportamentos discriminatórios (Carvalho et al., 2022), mesmo na ausência de experiência direta. O viés racial, portanto, abrange não apenas julgamentos avaliativos, mas também padrões observáveis de comportamento preferencial (por exemplo, preferência por promover um candidato branco em detrimento de um candidato negro com currículo idêntico) e ações discriminatórias

(por exemplo, atirar em uma pessoa negra desarmada enquanto se abstém de atirar em uma pessoa branca armada).

No entanto, o viés racial não é linear nem uniforme. Embora avaliações negativas possam aumentar a probabilidade de redução de preferências e de comportamentos discriminatórios, esses efeitos são fortemente dependentes do contexto. Em muitas situações, estímulos avaliados negativamente (por exemplo, “negro” pareado com “perigoso” ou “pouco inteligente”) são menos preferidos em contextos de seleção (como admissões escolares ou promoções no trabalho) ou mais frequentemente alvo de práticas excludentes (como abordagens policiais desproporcionais). Ao mesmo tempo, esses estímulos também podem evocar avaliações topograficamente positivas (por exemplo, “pessoas negras são fortes”). Contudo, tais avaliações positivas frequentemente coexistem com relações avaliativas negativas mais amplas e não necessariamente se traduzem em redução de desigualdades ou discriminação estrutural.

### **Viés Racial Explícito e Implícito**

O viés racial pode emergir tanto de respostas relacionais implícitas quanto explícitas. Do ponto de vista funcional, essa distinção não diz respeito à consciência ou intencionalidade, mas sim a se o comportamento é controlado diretamente por contingências ambientais ou indiretamente por redes relacionais previamente aprendidas (De Houwer & Moors, 2012) e pelas variáveis contextuais que influenciam o comportamento (Passarelli, Mizael et al., 2024; Watters et al., 2023). O viés explícito envolve respostas verbais ou manifestas que geralmente ocorrem quando os indivíduos dispõem de tempo suficiente para responder de acordo com contingências socialmente mediadas. Essas respostas são

comumente avaliadas por meio de instrumentos de autorrelato, como escalas do tipo Likert ou entrevistas estruturadas. Em contraste, o viés implícito consiste em respostas que são rápidas, que emergem sob pressão temporal ou em contextos ambíguos (Barnes-Holmes et al., 2010; Passarelli, Mizaél et al., 2024; Watters et al., 2023).

Nas últimas décadas, procedimentos computadorizados sob pressão de tempo, como o *Implicit Relational Assessment Procedure* (IRAP) e o *Function Acquisition Speed Test* (FAST), foram desenvolvidos para avaliar respostas relacionais rápidas e controladas pelo contexto. Ambas as tarefas utilizam janelas breves de resposta para reduzir a influência de estratégias de mascaramento social e revelar padrões avaliativos moldados pelas histórias de aprendizagem dos participantes.

O IRAP tem sido amplamente aplicado em pesquisas sobre viés racial como uma medida baseada em latência de respostas relacionais derivadas (Barnes-Holmes et al., 2010; Drake et al., 2015; Hussey et al., 2015; Power et al., 2017a; Power et al., 2017b). No Race-IRAP, os participantes completam blocos que se alinham ou contradizem redes relacionais culturalmente dominantes. Nos blocos consistentes, respostas como selecionar “Verdadeiro” para pareamentos Negro-negativo ou Branco-positivo são reforçadas; nos blocos inconsistentes, essa contingência é invertida. O principal índice, o escore d-IRAP, padroniza as diferenças de latência entre esses blocos. Valores positivos de d-IRAP, que refletem respostas mais lentas e menos precisas nos ensaios inconsistentes, indicam padrões relacionais pré-experimentais mais fortes. Por exemplo, Hussey et al. (2015) relataram escores d-IRAP significativos indicativos de viés racial

mesmo entre participantes que relataram atitudes igualitárias em medidas explícitas. Esses achados demonstram que restrições temporais podem revelar padrões relacionais que não emergem em condições que permitem mascaramento social ou controle reflexivo.

De forma semelhante, o FAST tem sido utilizado para examinar o viés racial avaliativo em diversos estudos recentes (Passarelli et al., 2024; Roche et al., 2025; Rodrigues et al., 2022). O FAST mede a rapidez com que os participantes aprendem relações arbitrárias estímulo–resposta sob pressão temporal. Os participantes são expostos a dois tipos de blocos: nos blocos consistentes, as respostas estão alinhadas com relações culturalmente dominantes ou experimentalmente estabelecidas; nos blocos inconsistentes, as contingências relacionais são invertidas. Respostas mais rápidas e precisas nos blocos consistentes indicam a força de redes relacionais pré-existentes<sup>2</sup>.

Rodrigues et al. (2022), por exemplo, utilizaram o *Function Acquisition Speed Test* (FAST) para examinar o viés racial avaliativo em redes relacionais verbais entre universitários. No FAST, apenas um estímulo é apresentado em cada tentativa, podendo ser uma face ou uma palavra, nunca pares de estímulos apresentados simultaneamente. O desempenho do participante é avaliado com

---

<sup>2</sup> O FAST mede a rapidez com que os participantes aprendem relações arbitrárias estímulo–resposta sob pressão temporal (O'Reilly et al., 2012; Roche et al., 2025). Os participantes são expostos a dois tipos de blocos: nos blocos consistentes, as respostas estão alinhadas com relações culturalmente dominantes, inferidas a partir de dados normativos e da literatura empírica sobre viés racial; nos blocos inconsistentes, as contingências relacionais são invertidas. Respostas mais rápidas e precisas nos blocos consistentes indicam a força de redes relacionais estabelecidas pela história de aprendizagem do indivíduo (Rodrigues et al., 2022; Watters et al., 2023), embora o procedimento não permita controle direto sobre a natureza ou extensão dessa história prévia.

base na rapidez e acurácia com que responde segundo contingências de classificação estabelecidas para cada bloco.

O procedimento incluiu dois blocos principais: um consistente com estereótipos raciais e outro inconsistente com estereótipos. No bloco consistente, os participantes eram instruídos a pressionar a mesma tecla (por exemplo, “Z”) tanto para faces negras quanto para palavras negativas, e outra tecla (por exemplo, “M”) tanto para faces brancas quanto para palavras positivas. No bloco inconsistente, essas contingências eram invertidas: a mesma tecla era utilizada para faces negras e palavras positivas, enquanto a outra tecla era utilizada para faces brancas e palavras negativas. Os participantes responderam de forma mais rápida e precisa nos blocos consistentes com o estereótipo racial. Além disso, os efeitos observados no FAST não corresponderam aos autorrelatos explícitos de viés racial. Essa dissociação sugere que tarefas relacionais sob pressão temporal, como o FAST, são sensíveis à emissão de respostas relacionais breves e automatizadas, menos influenciadas por estratégias de autorrelato socialmente mediadas.

De modo geral, avaliações implícitas medidas por tarefas baseadas em latência refletem padrões relacionais aprendidos que podem exercer forte controle sobre o comportamento em ambientes naturais. Em situações cotidianas que envolvem pressão de tempo, ambiguidade ou baixo monitoramento social, essas redes relacionais podem orientar ações como atenção, esquiva e tomada de decisão, mesmo quando os relatos explícitos são igualitários. Assim, o viés implícito é mais bem compreendido não como um construto mental oculto, mas como um padrão de responder relacional moldado pela história de aprendizagem,

capaz de influenciar resultados discriminatórios em situações realísticas quando variáveis contextuais evocam ou fortalecem esses padrões.

Portanto, a presente revisão sistemática surge neste momento crítico com o objetivo de sintetizar achados empíricos sobre o viés racial antinegro no âmbito da pesquisa em análise do comportamento. Examinamos como a análise do comportamento contemporânea compreende, mensura e intervém sobre o viés antinegro a partir de seu arcabouço teórico próprio. Ao conectar os fundamentos históricos da análise do comportamento ao seu papel emergente na pesquisa sobre viés racial, esta revisão busca contribuir para o avanço de uma ciência comportamental empiricamente fundamentada, socialmente responsiva e capaz de enfrentar as injustiças raciais.

### **Método**

Foi conduzida uma revisão sistemática de acordo com as diretrizes PRISMA (Page et al., 2021). As buscas foram realizadas nas bases APA PsycINFO e APA PsycArticles e foram restritas a 16 periódicos historicamente ou conceitualmente alinhados à análise do comportamento (ver Apêndice I). Os termos de busca incluíram “racial bias”, “racism” e “racial prejudice”. Os estudos foram elegíveis para inclusão se fossem empíricos (experimentais, quase-experimentais, correlacionais, de levantamento, transversais ou longitudinais), envolvessem participantes humanos e apresentassem dados originais que examinassem especificamente o viés racial ou racismo ou preconceito racial antinegro. Esse critério foi atendido quando o estudo mensurava, analisava ou intervinha sobre respostas relacionais ou avaliativas dirigidas a pessoas negras, comparações entre pessoas negras e brancas, ou estímulos racialmente

relevantes para os quais resultados específicos relacionados a pessoas negras fossem relatados separadamente.

Para inclusão, os estudos também precisavam demonstrar fundamentação comportamental ou contextual-comportamental, operacionalizada pelo uso de conceitos da análise do comportamento (por exemplo, reforçamento, responder relacional derivado, transformação de funções de estímulo, controle contextual), procedimentos analítico-comportamentais (por exemplo, IRAP, FAST, treino baseado em equivalência, MTS, reforçamento diferencial) ou intervenções inseridas na Ciência Comportamental Contextual, incluindo ACT e outros modelos funcional-analíticos. Apenas artigos publicados em periódicos revisados por pares foram considerados.

Foram excluídos estudos de natureza teórica, conceitual, editorial ou de revisão, bem como aqueles que abordavam o racismo de forma geral sem fornecer análises específicas sobre o viés antinegro. Pesquisas baseadas exclusivamente em modelos cognitivos ou sociocognitivos, sem interpretação funcional ou comportamental, também foram excluídas, assim como estudos que examinavam preconceito direcionado a outros grupos, a menos que dados específicos sobre pessoas negras fossem apresentados separadamente. Estudos com organismos não humanos ou que avaliassem fenômenos não raciais também foram excluídos.

O processo de revisão seguiu quatro etapas: (1) realização das buscas nas bases de dados utilizando palavras-chave, com aplicação de filtros para restringir os resultados aos periódicos previamente definidos; (2) exportação dos resultados para a plataforma Rayyan para revisão sistemática, com triagem de títulos,

resumos e palavras-chave com base nos critérios de exclusão; (3) leitura dos textos completos para confirmar a elegibilidade conforme os critérios de inclusão; e (4) extração e categorização dos dados em três grupos: (a) estudos descritivos, longitudinais ou de levantamento que mensuraram a relação entre questões raciais e construtos psicológicos (por exemplo, inflexibilidade, depressão); (b) estudos experimentais ou quase-experimentais sobre medidas de viés racial; e (c) estudos experimentais ou quase-experimentais de intervenção. Os dados extraídos incluíram autor/ano, país, delineamento do estudo, características dos participantes, medidas e intervenções.

### **Resultados**

A busca realizada com base nas palavras-chave, filtros e periódicos-alvo (ver Apêndice I) identificou 84 artigos publicados entre 2004 e 2024. Durante a fase de triagem, 55 artigos foram excluídos com base nos critérios de exclusão previamente definidos, por se tratar de trabalhos teóricos, revisões ou meta-análises. Em seguida, os textos completos dos 30 artigos remanescentes foram examinados de acordo com os critérios de inclusão, e 10 foram excluídos por não abordarem diretamente o viés racial antinegro, resultando na inclusão de 20 estudos empíricos.

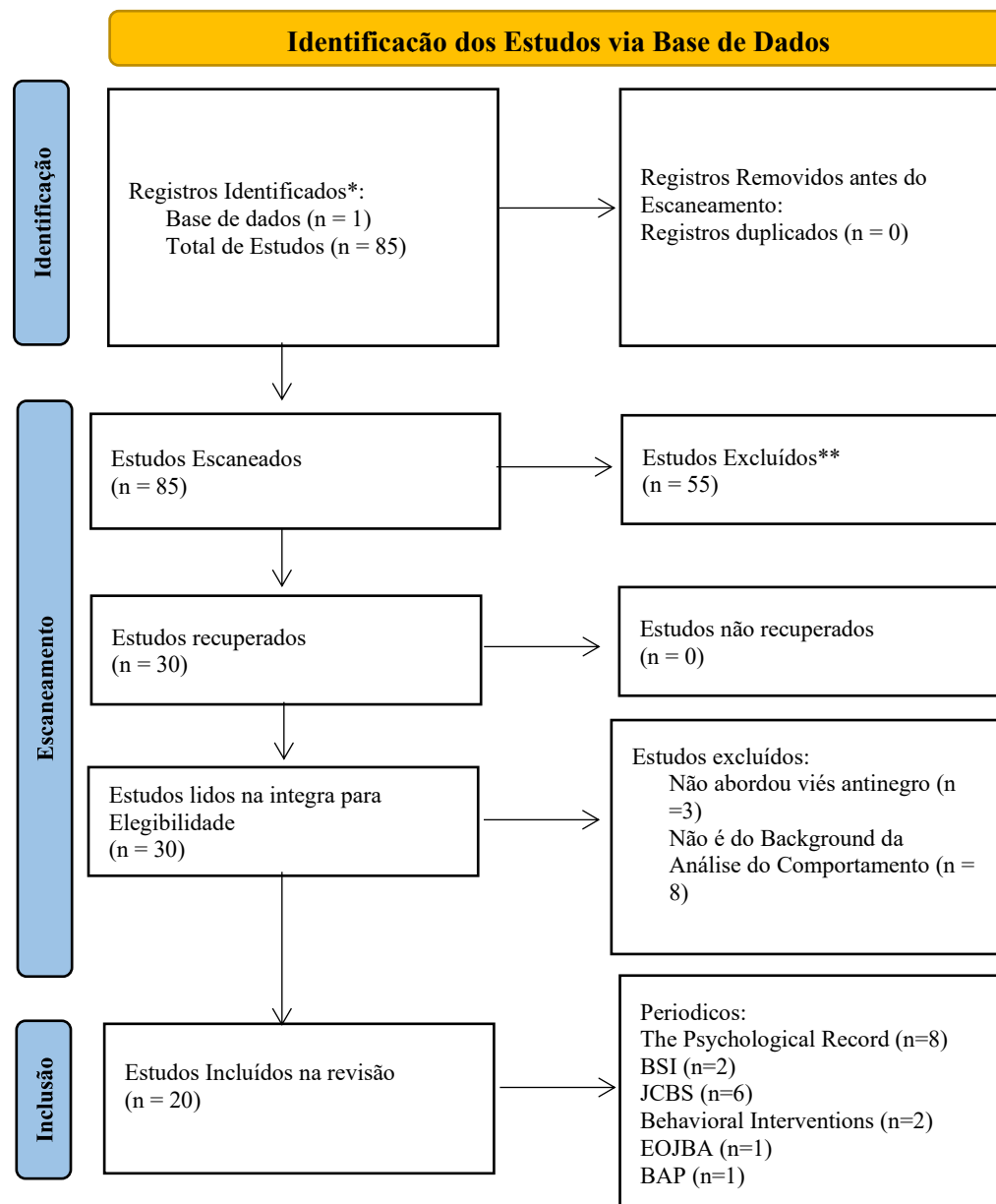
Esses estudos abrangem o período de 2007 a 2025, iniciando com uma única publicação em 2007, seguida por três estudos entre 2010 e 2014 e dois em 2015. A frequência de publicações aumentou posteriormente, com cinco estudos publicados entre 2016 e 2020 e um crescimento acentuado entre 2022 e 2024, período que concentrou sete estudos e se configurou como o mais produtivo, distribuídos em três publicações em 2022, duas em 2023 e duas em 2024, além

de um estudo adicional publicado em 2025. No total, aproximadamente 63% dos estudos foram publicados nos últimos cinco anos, indicando uma clara tendência de crescimento da atividade de pesquisa nessa área.

Quanto aos periódicos de publicação, sete estudos foram publicados no *The Psychological Record*, seis no *Journal of Contextual Behavioral Science* (JCBS), dois em *Behavior and Social Issues* (BSI), dois em *Behavioral Interventions* e um em cada um dos seguintes periódicos: *European Journal of Behavior Analysis* (EJoBA) e *Behavior Analysis in Practice* (BAP). A Figura 1 apresenta o diagrama de fluxo PRISMA, detalhando procedimento de busca.

**Figura 1.**

*Diagrama PRISMA com procedimento de busca e resultados.*



### Características Metodológicas dos Estudos

Os estudos selecionados abrangeram uma ampla gama de características metodológicas, incluindo diversidade de participantes, delineamentos de pesquisa

e instrumentos de medida (ver Tabela 1). Os estudos incluídos foram conduzidos exclusivamente em três países, sendo 11 originários dos Estados Unidos, cinco do Brasil e quatro da Irlanda. Os estudos utilizaram amostras heterogêneas, incluindo indivíduos que apresentavam viés racial, participantes negros expostos a experiências de racismo ou microagressões, profissionais da Análise do Comportamento Aplicada (ABA) e crianças. Dos 20 estudos incluídos, quatro apresentaram delineamentos quantitativos descritivos (estudos transversais ou correlacionais), enquanto os 15 restantes empregaram metodologias experimentais. Entre os estudos experimentais, dois utilizaram delineamentos de multielementos, sete adotaram delineamentos intra-sujeitos, três empregaram delineamentos mistos (entre e intra-sujeitos), um foi um ensaio clínico randomizado (RCT) e um consistiu em um experimento de grupos contrabalanceados.

De modo geral, a avaliação metodológica revelou um padrão consistente entre os vinte estudos. Entre os delineamentos quantitativos descritivos (Catagnus et al., 2022; Graham et al., 2015; Florez et al., 2019; LoPresti et al., 2022), todos apresentaram questões de pesquisa claras e estratégias de mensuração adequadas; entretanto, as amostras foram não representativas e nenhum estudo descreveu de forma suficiente o fluxo de recrutamento para permitir a avaliação de viés de não resposta. O ensaio clínico randomizado (West et al., 2013) descreveu o procedimento de randomização, mas não relatou equivalência basal entre os grupos nem procedimentos de cegamento, o que limita a validade interna.

Os experimentos quantitativos não randomizados, incluindo intervenções baseadas em ACT (Lillis & Hayes, 2007; Banks et al., 2021; Williams et al., 2020),

estudos laboratoriais comportamentais de mensuração (Barnes-Holmes et al., 2010; Power et al., 2017; Belisle et al., 2023; Roche et al., 2025; Passarelli, Mizael et al., 2024), treinamentos relacionais baseados em equivalência com crianças (de Carvalho & de Rose, 2014; Mizael et al., 2016; Mizael et al., 2021; Passarelli et al., 2024), experimentos de enquadramento relacional (Suárez et al., 2024) e procedimentos de reforçamento diferencial com feedback (Romani et al., 2024), demonstraram forte clareza procedimental e medidas de desfecho apropriadas. No entanto, todos dependeram de amostras pequenas e de conveniência, e nenhum controlou analiticamente possíveis variáveis de confusão.

Os instrumentos de autorrelato foram agrupados em três categorias. A primeira incluiu medidas de sintomas psicológicos, afeto e repertório experiencial, tais como ansiedade, estresse, humor, esquiva experiencial, imagética, sofrimento emocional, sintomas de trauma e autocompaixão (utilizadas em Banks et al., 2021; West et al., 2013; Florez et al., 2019; LoPresti et al., 2022; Williams et al., 2024; Graham et al., 2015; Catagnus et al., 2022). A segunda categoria incluiu medidas de atitudes raciais e experiências relacionadas à raça, incluindo cegueira racial (*colorblindness*), racismo moderno, identidade étnica, microagressões, histórico de discriminação e tendências microagressivas (utilizadas em Banks et al., 2021; West et al., 2013; Florez et al., 2019; LoPresti et al., 2022; Williams et al., 2024; Lillis & Hayes, 2007; Catagnus et al., 2022; Barnes-Holmes et al., 2010; Power et al., 2017; Mizael et al., 2016, 2020; Roche et al., 2025). A terceira categoria incluiu medidas avaliativas e baseadas em valência, tais como diferenciais semânticos, a escala de valência emocional, escalas do tipo Likert, termômetros afetivos e avaliações modificadas baseadas em equivalência (utilizadas em Passarelli et al.,

2024; Williams et al., 2024; Romani et al., 2023; Belisle et al., 2023; Barnes-Holmes et al., 2010; Power et al., 2017; Mizael et al., 2016, 2020; Suárez et al., 2024; Roche et al., 2025).

Embora diversos, a maioria dos instrumentos teve origem na psicologia cognitiva. Menos da metade dos (8 de 20) estudos incorporou avaliações comportamentais ou relacionais, tais como o IRAP, Testes de Equivalência Modificados, Testes Relacionais ou o FAST (Barnes-Holmes et al., 2010; Power et al., 2017; Mizael et al., 2016, 2020; Passarelli et al., 2024; Suárez et al., 2024; Belisle et al., 2023; Roche et al., 2025).

As intervenções também variaram amplamente, incluindo exercícios de clarificação de valores, conteúdos psicoeducativos sobre racismo e treinamentos baseados na Terapia de Aceitação e Compromisso (ACT), em Estratégias Cognitivo-Comportamentais (CBS), na Psicoterapia Analítico-Funcional (FAP), em Treino de Discriminação Condicional (por exemplo, *Matching-to-Sample*) e em Reforçamento Diferencial com Feedback (Banks et al., 2021; LoPresti et al., 2022). Essas estratégias tiveram como objetivo modificar repertórios de enfrentamento emocional, aumentar respostas de *mindfulness*<sup>3</sup> e empatia e alterar o responder relacional avaliativo direcionado a estímulos raciais (Romani et al., 2024; Williams et al., 2020). No entanto, apenas cinco estudos incluíram avaliações de seguimento (*follow-up*) (Lillis & Hayes, 2007; Williams et al., 2020; Banks et al.,

---

<sup>3</sup> *Mindfulness* refere-se a um estado de atenção plena e deliberada ao momento presente, caracterizado pela observação das experiências internas e externas, como pensamentos, emoções e sensações corporais, sem julgamento e com abertura. No contexto psicológico, práticas de *mindfulness* têm sido utilizadas para ampliar a consciência do próprio comportamento e de seus determinantes contextuais, favorecendo maior flexibilidade comportamental e redução de respostas automáticas governadas por histórias prévias rígidas.

2021; Mizael et al., 2021; Passarelli et al., 2024), e seis estudos testaram a generalização (Carvalho & de Rose, 2014; Mizael et al., 2016; Mizael et al., 2021; Passarelli et al., 2024; Suárez et al., 2024).

### Tabela 1.

#### *Características Metodológicas dos Estudos Selecionados*

<b>Autores</b>	<b>País</b>	<b>Participantes</b>	<b>Delineamento</b>	<b>Medidas</b>	<b>Intervenção</b>
<b>Lillis &amp; Hayes (2007)</b>	EUA	32 estudantes de graduação (13 homens, 20 a 37 anos)	Delineamento experimental entre grupos com contrabalanceamento	Intenções comportamentais autorrelatadas, mudanças relacionadas ao modelo ACT	Grupo 1: sessão baseada em Terapia de Aceitação e Compromisso; Grupo 2: palestra educativa sobre diferenças raciais. Avaliação de follow-up realizada 7 dias depois.
<b>Barnes-Holmes et al. (2010)</b>	Irlanda	31 irlandeses brancos (15 homens, 16 mulheres, M = 21 anos)	Delineamento misto (between-subjects × within-subjects)	Medidas de autorrelato: Discrimination and Diversity Scale (DDS), Modified Modern Racism Scale (MMRS), Likert Scale e Implicit Relational Assessment Procedure (IRAP)	Não
<b>West et al. (2013)</b>	EUA	14 participantes (2 homens e 12 mulheres) autoidentificados como Black American	Ensaio clínico randomizado (Randomized Controlled Trial)	Schedule of Racist Events (SRE), Vividness of Visual Imagery Questionnaire (VVIQ), Positive and Negative Affect Schedule (PANAS), Subjective Units of Distress Scale (SUDS)	Psicoeducação sobre racismo e opressão racial internalizada, associada a estratégias baseadas em Terapia de Aceitação e Compromisso.

<b>Carvalho &amp; de Rose (2014)</b>	Brasil	4 crianças, 7 a 10 anos (com viés racial)	Delineamento experimental intra-sujeitos (within-subjects)	Modified Equivalence Test (AC3) e Likert Scale	Treino de Emparelhamento de Acordo com o Modelo ( <i>Matching-to-Sample</i> ).
<b>Drake et al. (2015)</b>	EUA	57 estudantes de graduação em Psicologia (39 homens, M = 19,1 anos)	Delineamento intra-sujeitos com medidas repetidas	Questionário demográfico, MMRS, Semantic Differential Scale (SDS), Social Dominance Orientation (SDO)	Não
<b>Graham, West &amp; Roemer (2015)</b>	EUA	57 estudantes afro-americanos de graduação e pós-graduação (M = 28,89; 74% mulheres)	Delineamento correlacional, quantitativo descritivo	Schedule of Racist Events – Past Year (SRE-R), Valued Living Questionnaire (VLQ), Depression Anxiety Stress Scales (DASS-21)	Não
<b>Mizael et al. (2016)</b>	Brasil	13 crianças, 8 a 10 anos (com viés racial)	Delineamento experimental intra-sujeitos (within-subjects)	Self-Assessment Manikin (SAM), AC3 e IRAP	Treino de Matching-to-Sample
<b>Power et al. (2017a)</b>	Irlanda	16 indivíduos brancos residentes na Irlanda (8 mulheres, M = 25 anos)	Delineamento experimental intra-sujeitos	IRAP e registros de Electroencephalography (EEG)	Não
<b>Power et al. (2017b)</b>	Irlanda	22 participantes negros (11 mulheres, M = 22 anos)	Delineamento misto (entre-sujeitos × intra-sujeitos)	IRAP, Discrimination and Diversity Scale (DDS), Semantic Differential Scale (SDS) e Feeling Thermometer	Não
<b>Florez et al. (2019)</b>	EUA	253 estudantes universitários brancos (M = 19,06; DP = 2,06; 197 mulheres)	Estudo correlacional transversal	Meaning in Life Questionnaire–Presence (MLQ-P), Prejudice Against Blacks scale (PAB), Portrait Values Questionnaire–Revised Revised–Short Term (PVQ-RR-ST), Valuing Questionnaire–Obstruction (VQ-O)	Não
<b>Mizael et al. (2020)</b>	Brasil	46 crianças de 8 a 10 anos (11 negras, 27 meninas), com viés racial demonstrado no pré-teste	Delineamento misto (entre-grupos × intra-sujeitos)	Self-Assessment Manikin (SAM), Modified Equivalence Test e IRAP	Treino de Emparelhamento de Acordo com o Modelo ( <i>Matching-to-Sample</i> ). Avaliação de follow-up

					realizada após 6 semanas
<b>Williams et al. (2020)</b>	EUA	44 estudantes universitários (31 mulheres; 24 brancos e 20 negros)	Delineamento misto (entre-grupos × intra-sujeitos)	PANAS, Allophilia General/Workshop, Intergroup Anxiety and Avoidance (IAA-6), Satisfaction. Para participantes negros: Multigroup Ethnic Identity Measure (MEIM-12); para participantes brancos: Color-Blind Racial Attitudes Scale (CoBRAS), Symbolic Racism Scale (SR2K), Color-Conscious Awareness Scale (CCAS – Attitudes and Behaviors)	Grupo 1: Workshop de Harmonia Racial ( <i>Racial Harmony Workshop -RHW</i> ); Grupo 2: Controle. Avaliação de follow-up realizada após 4 semanas
<b>Banks et al. (2021)</b>	EUA	20 mulheres negras, 18–60 anos (M = 36,26; DP = 14,92)	Delineamento experimental intra-sujeitos com medidas repetidas	DASS-21, IROS, AAQ-II, ISS, SBWS, SRE, VVIQ, PANAS, SUDS	Exercício de clarificação de valores ( <i>values clarification</i> ). Avaliação de follow-up realizada após 3 meses
<b>LoPresti et al. (2022)</b>	EUA	202 participantes negros (idade média = 25,43; 132 mulheres)	Estudo correlacional transversal	Demographic Questionnaire, UMass Boston Comprehensive Race Implicit Bias Measures, escalas de experiências raciais e saúde psicológica	Não
<b>Catagnus et al. (2022)</b>	EUA	152 participantes (117 mulheres, M = 37,4 anos)	Estudo qualitativo por survey	Questionário sobre experiência emocional e racismo	Não
<b>Romani et al. (2023)</b>	EUA	26 adultos (21 mulheres, 5 homens), idades entre 27 e 56 anos (M = 38), empregados em um hospital infantil	Delineamento multielemento com linha de base e duas condições de treino (Grupo 1 e Grupo 2)	Questionários contendo 15 itens cada: 5 microagressões, 5 não-exemplos e 5 itens controle	Reforço diferencial com <i>feedback</i> corretivo para auxiliar profissionais de saúde a discriminar exemplos e não-exemplos de microagressões raciais
<b>Belisle et al. (2023)</b>	EUA	34 participantes	Estudo correlacional transversal	Medidas de Resposta Relacional e Características do Estímulo	Não

		(M = 23 anos; 31 mulheres)			
<b>Passarelli et al. (2024)</b>	Brasil	9 participantes (M = 8,22 anos; faixa etária 8–9 anos; 5 meninos e 4 meninas)	Delineamento experimental intra-sujeitos com medidas repetidas	Relational Test, Function Acquisition Speed Test (FAST), Likert Scale, Doll Test	Treino de Emparelhamento de Acordo com o Modelo ( <i>Matching-to-Sample</i> ). Avaliação de follow-up realizada após 6 semanas
<b>Passarelli et al. (2024)</b>	Brasil	34 participantes (M = 8,6 anos; 24 meninas)	Delineamento experimental intra-sujeitos com medidas repetidas	Likert Scale e Function Acquisition Speed Test (FAST)	Não
<b>Suárez et al. (2024)</b>	EUA	2 participantes (mulheres, 31 e 39 anos)	Delineamento experimental de caso único multielemento com controle sem tratamento	Questionário de empatia direcionada a diferentes grupos raciais	Treino relacional com molduras de coordenação e distinção
<b>Roche, Passarelli &amp; Wall (2025)</b>	Irlanda	27 participantes brancos e 25 não brancos	Delineamento misto (entre-sujeitos × intra-sujeitos)	FAST, Discrimination and Diversity Scale (DDS), Modified Modern Racism Scale (MMRS)	Não

*Nota.* O Apêndice II apresenta as traduções referentes aos nomes de escalas. As traduções a seguir são livres e descritivas, apresentadas apenas para fins de compreensão geral do leitor

## Estudos Transversais e Correlacionais

Quatro estudos de levantamento transversal ou correlacional (Catagnus et al., 2022; Florez et al., 2019; Graham et al., 2015; LoPresti et al., 2022) investigaram fatores psicológicos subjacentes às atitudes raciais, à saúde mental e aos valores. Esses estudos examinaram relações entre flexibilidade psicológica, regulação emocional, racismo percebido, vida orientada por valores, preconceito, sofrimento psicológico e estratégias de enfrentamento.

Dois estudos examinaram os efeitos do racismo a partir da perspectiva dos indivíduos que o vivenciam diretamente, utilizando medidas voltadas à

caracterização de experiências racistas e de seus impactos emocionais e psicológicos. Os instrumentos empregados capturam tanto eventos racistas explícitos, como comentários ofensivos, xingamentos e exclusões fenotipicamente motivadas, quanto formas mais sutis de violência racial, como microagressões, invisibilização e inferiorização não declarada. Essa distinção é relevante porque evidências indicam que diferentes formas de violência racial podem produzir impactos psicológicos distintos, com microagressões crônicas associadas a cargas traumáticas acumulativas mesmo na ausência de eventos explícitos (LoPresti et al., 2022).

Graham et al. (2015) examinaram se a vida orientada por valores modera o impacto emocional de experiências racistas em adultos afro-americanos. Cinquenta e sete estudantes de graduação e pós-graduação ( $M = 28,89$ ; 74,2% mulheres) responderam a medidas de eventos racistas (Schedule of Racist Events – Past Year; SRE-R), vida orientada por valores (Valued Living Questionnaire; VLQ) e sintomas de depressão, ativação ansiosa e ansiedade geral (Depression Anxiety Stress Scales – 21 itens; DASS-21). A vida orientada por valores apresentou correlações negativas significativas com todos os conjuntos de sintomas. As análises de moderação indicaram que experiências racistas predisseram maior sofrimento emocional apenas entre participantes com baixos níveis de vida orientada por valores, sugerindo que o engajamento em domínios de vida pessoalmente significativos pode atenuar os efeitos psicológicos do racismo.

LoPresti et al. (2022) exploraram as relações entre microagressões raciais, sintomas de trauma e autocompaixão em 202 adultos negros norte-americanos

recrutados online e por meio de redes universitárias. Os participantes responderam a medidas de microagressões raciais (Racial Microaggressions Scale; RMAS), sintomas de estresse pós-traumático (Posttraumatic Stress Disorder Checklist for DSM-5; PCL-5) e autocompaixão (Self-Compassion Scale – Short Form; SCS-SF). As microagressões apresentaram correlações fortes e positivas com os sintomas de estresse pós-traumático, com aproximadamente 70% dos participantes pontuando acima do ponto de corte clínico para transtorno de estresse pós-traumático. A autocompaixão moderou apenas os efeitos das microagressões relacionadas à invisibilidade, caracterizadas por práticas sociais sutis de omissão, silenciamento ou não reconhecimento da identidade e das experiências do indivíduo, que operam como contingências aversivas por ausência de validação social. Essa moderação reduziu, mas não eliminou, o impacto dessas experiências sobre os sintomas traumáticos. Os achados destacam o elevado ônus psicológico associado à exposição crônica a microagressões e o papel protetivo parcial da autocompaixão.

Já o estudo de Florez et al. (2019) abordou a temática a partir da perspectiva dos perpetradores do racismo, isto é, de indivíduos que emitem respostas preconceituosas. Os autores investigaram se o sentido da vida prediz o preconceito racial e se essa relação é mediada por autotranscendência e inflexibilidade psicológica, examinando fatores associados à probabilidade de engajamento em comportamentos preconceituosos. A amostra foi composta por 253 estudantes universitários brancos (M = 19,06; 77,9% mulheres), que responderam a medidas validadas de sentido da vida (MLQ-P), preconceito antinegro (PAB), valores de autotranscendência (PVQ-RR-ST) e inflexibilidade

psicológica (VQ). Maiores níveis percebidos de sentido da vida estiveram associados a menores níveis de preconceito racial. As análises de mediação indicaram que autotranscendência e inflexibilidade psicológica explicaram conjuntamente parte dessa relação. Do ponto de vista analítico-comportamental, esses construtos podem ser reinterpretados como repertórios verbais sob controle de contingências específicas: a autotranscendência como responder governado por regras orientadas a valores coletivos, e a inflexibilidade psicológica como responder sob controle rígido de estímulos verbais aversivos, ambos passíveis de investigação experimental direta.

Uma das pesquisas voltou-se ao contexto da atuação profissional, investigando como analistas do comportamento experienciam emoções no trabalho clínico e como percebem emoções associadas ao racismo. Catagnus et al. (2022) conduziram uma pesquisa quantitativa-descritiva para examinar como profissionais da análise do comportamento vivenciam emoções em seu trabalho clínico e como percebem emoções relacionadas ao racismo. Utilizando uma estratégia de amostragem em bola de neve por meio de listas de e-mail e redes sociais, o estudo reuniu respostas de 152 profissionais, em sua maioria mulheres e analistas do comportamento certificados (BCBAs). O questionário incluiu itens fechados e abertos elaborados pelos autores para avaliar conforto com a expressão emocional, reações a situações emocionalmente carregadas e percepções sobre discutir raça em contextos profissionais. Os resultados revelaram grande variabilidade: muitos participantes relataram sentir-se confortáveis ao discutir emoções com clientes ou colegas, enquanto uma parcela significativa relatou desconforto ao expressar ou presenciar comportamentos

emocionais intensos. Aproximadamente um quarto dos participantes indicou desconforto ao discutir questões relacionadas à raça, embora a grande maioria tenha concordado que compreender emoções relacionadas ao racismo é importante para a prática profissional. A maioria também endossou a ideia de que a análise do comportamento aplicada deve abordar o racismo e desenvolver intervenções voltadas a emoções associadas à opressão.

### **Estudos de Medidas Comportamentais de Identificação de Preconceito Racial**

A maioria dos estudos comportamentais sobre viés racial concentrou-se na validação de medidas implícitas, sendo Belisle et al. (2023) a principal exceção ao empregar paradigmas comportamentais alternativos. Barnes-Holmes et al. (2010), por exemplo, conduziram um estudo experimental examinando os efeitos do contexto privado versus público e da latência de resposta sobre a estereotipia pró-branca e anti-negra em 31 participantes irlandeses brancos (15 homens, 16 mulheres,  $M = 21$  anos). O estudo utilizou o IRAP para comparar o desempenho em blocos que exigiam respostas pró-branco/antinegro ou pró-negro/antibranco, além de manipular se os participantes acreditavam que suas respostas eram públicas ou privadas, bem como o critério de latência exigido para respostas corretas. Os resultados mostraram que o contexto público não reduziu o viés pró-branco e que critérios de latência mais curtos produziram respostas mais fortemente consistentes com os estereótipos, indicando que o responder relacional imediato exerceu maior controle sobre o desempenho do que pistas contextuais relacionadas à avaliação pública. Esse padrão contrasta com o observado em medidas explícitas, nas quais a audiência tende a exercer controle mais

pronunciado, como demonstrado por Rogers e Prentice-Dunn (1981), em que a avaliação pública alterou a magnitude dos choques simulados (fictícios) que os participantes acreditavam estar administrando a participantes negros.

Drake et al. (2015) conduziram um estudo experimental quantitativo examinando a confiabilidade e a validade convergente de um IRAP racial administrado duas vezes em sucessão a uma amostra diversa de estudantes universitários (N = 57, com subamostras equilibradas de participantes negros e brancos). Os participantes completaram medidas de autorrelato, tais como Escala de Racismo Moderno, Escala de Diferencial Semântico e Orientação para Dominância Social, seguidas da aplicação do IRAP-Racial. Cabe destacar que o IRAP era formado por dois blocos, um bloco consistente com estereótipos raciais antinegro e pró-branco e um bloco inconsistente com esses estereótipos, resultando em pró-negro e antibranco. Os resultados indicaram confiabilidade interna média, confiabilidade teste–reteste de pequena a média magnitude e correlações consistentes de magnitude moderada a alta com escores de racismo moderno, sustentando a validade convergente. A análise por tipo de tentativa revelou padrões sutis de favoritismo intragrupal (pró-negro entre participantes negros; pró-branco entre participantes brancos), em vez de desvalorização explícita do grupo externo, destacando uma especificidade funcional não capturada por métricas como o IAT<sup>4</sup>. Contudo, a taxa de evasão foi elevada e

---

<sup>4</sup> O *Implicit Association Test* (IAT) é uma tarefa computadorizada que avalia associações implícitas entre categorias (por exemplo, grupos sociais) e atributos avaliativos (positivos ou negativos), com base na velocidade das respostas em classificações sob restrição temporal. Diferenças no desempenho entre blocos consistentes e inconsistentes são interpretadas como indicadores da força relativa dessas associações, menos dependentes de autorrelato consciente.

assimétrica racialmente (especialmente alta entre participantes negros), o que limita a estabilidade dos resultados e a interpretação das comparações entre subgrupos.

Power et al. (2017a) conduziram um estudo combinando eletroencefalografia e IRAP (EEG–IRAP) com 16 adultos irlandeses brancos (M = 25). A qualidade metodológica foi considerada moderada, com medidas e análises apropriadas; entretanto, a representatividade foi limitada pelo tamanho reduzido e pela homogeneidade racial da amostra, e o viés de não resposta permaneceu pouco claro devido a exclusões substanciais relacionadas ao EEG ( $\approx 43\%$ ), sem avaliação de se esses critérios introduziram viés sistemático. Os dados comportamentais revelaram viés pró-branco/antinegro, corroborado por padrões de potenciais relacionados a eventos (ERPs) que sugeriram maior conflito cognitivo em blocos incongruentes com os estereótipos.

Outro estudo de Power et al. (2017b) examinou viés racial entre participantes negros e brancos utilizando o IRAP e medidas explícitas. A amostragem foi adequada, mas não representativa, uma vez que a pequena subamostra de participantes negros (todos imigrantes nigerianos) e a subamostra branca (estudantes nascidos na Irlanda) não refletiam a população irlandesa mais ampla. Importante destacar que este estudo apresentou a evidência mais clara de viés de não resposta: vários participantes negros foram excluídos por não atenderem aos critérios do IRAP, enquanto dados comparáveis de exclusão para participantes brancos não foram reportados. Os resultados revelaram padrões de favoritismo intragrupal e demonstraram que o IRAP apresentou maior validade preditiva do que as escalas explícitas.

Pesquisas comportamentais mais recentes ampliaram o foco para além do IRAP. Passarelli, Mizaël et al. (2024) conduziram um estudo de medidas repetidas com 34 crianças ( $M = 8,6$ ), avaliando viés racial ao longo de três sessões por meio de duas medidas comportamentais: o FAST e avaliações explícitas do tipo Likert para avaliar valência emocional das faces. Ao longo das sessões, um subconjunto de 11 crianças apresentou viés implícito antinegro estável no FAST, enquanto as avaliações explícitas não indicaram preferência racial, evidenciando uma dissociação entre responder avaliativo implícito e explícito na infância.

Roche et al. (2025) avaliaram a sensibilidade do FAST para detectar viés implícito antinegro. Utilizando uma amostra online, os participantes completaram o FAST juntamente com medidas explícitas de atitudes raciais. O FAST detectou de forma consistente viés implícito antinegro, enquanto as escalas de autorrelato não indicaram preferência racial. Essa dissociação replicou o padrão amplamente documentado segundo o qual tarefas relacionais rápidas, sob pressão temporal, revelam padrões avaliativos que não são capturados por medidas explícitas.

Belisle et al. (2023) utilizaram a Teoria da Densidade Relacional, uma extensão quantitativa da RFT que modela como redes de respostas relacionais torna-se progressivamente mais densas, coesas e resistentes à mudança, para analisar a organização relacional subjacente ao responder preconceituoso. Os participantes completaram tarefas que avaliaram tendências de aproximação–evitação e o agrupamento relacional de estímulos raciais. Os resultados indicaram que redes relacionais envolvendo rostos negros tendiam a se agrupar com funções de evitação, enquanto redes relacionadas a estímulos brancos se

agruparam com funções de aproximação, revelando padrões estáveis consistentes com viés antinegro.

### **Estudos de Intervenções no Viés Racial Antinegro**

Exatamente metade dos estudos selecionados (10 de 20) implementou intervenções com o objetivo de reduzir o viés racial ou mitigar o sofrimento relacionado ao racismo. Desses, quatro intervenções foram direcionadas a crianças e seis a adultos. Para apresentar esses achados de forma clara, a literatura de intervenção foi organizada em três categorias metodológicas. O primeiro grupo compreende intervenções voltadas para crianças, baseadas em treino relacional por equivalência. O segundo inclui estudos laboratoriais com adultos que manipularam o controle contextual por meio de enquadramento relacional ou reforçamento diferencial. O terceiro grupo abrange intervenções baseadas em Terapia de Aceitação e Compromisso (ACT; *Acceptance and Commitment Therapy*) e em processos baseados em valores, implementadas com adultos.

#### *Treino de Equivalência de Estímulos com Crianças*

Um conjunto de quatro estudos (de Carvalho & de Rose, 2014; Mizaél et al., 2016; Mizaél et al., 2021; Passarelli et al., 2024) empregou procedimentos de equivalência de estímulos para alterar as funções avaliativas de estímulos raciais em crianças. Todos os estudos seguiram a mesma estratégia básica. Inicialmente, as crianças foram triadas por meio de uma tarefa de matching-to-sample (MTS; emparelhamento ao modelo), que identificou aquelas que tendiam a selecionar faces negras na presença do símbolo negativo a rostos negros, indicando viés pré-experimental. Em seguida, todos os estudos implementaram treinos de

discriminação condicional AB e BC, projetados para produzir relações derivadas AC e CA sem treino direto. Tipicamente, A correspondia a um símbolo positivo, B a um estímulo abstrato neutro e C a um rosto negro. O objetivo era criar uma classe de equivalência na qual a função avaliativa positiva de A (símbolo positivo) fosse transferida para C (rosto negro). Embora a estrutura básica tenha sido compartilhada entre os quatro estudos, cada um introduziu variações procedimentais específicas destinadas a melhorar a formação das classes, fortalecer a manutenção ou testar a generalização.

De Carvalho e de Rose (2014) implementaram o treino AB–BC com quatro crianças. Embora todas tenham aprendido as relações treinadas, apenas uma formou equivalência entre símbolos positivos (A1) e rostos negros (C1), sem evidência de transferência de função. As três crianças restantes mantiveram ou até intensificaram suas avaliações negativas originais, sugerindo que redes relacionais pré-existentes exerceram forte controle contextual, limitando a emergência da nova classe de equivalência.

Mizael et al. (2016) ampliaram o procedimento ao incorporar parâmetros que são considerados facilitadores da formação de classe de equivalência, como o sequenciamento do simples para o complexo (simple-to-complex), treino misto (mixed training), além de condições controladas de feedback e atraso, comparando também MTS simultâneo e atrasado. Esses refinamentos produziram resultados substancialmente melhores: todas as 13 crianças formaram as classes-alvo, e nove mantiveram a relação Negro–positivo mesmo seis semanas após a intervenção. As avaliações explícitas medidas pela escala SAM mostraram que,

antes do treino, rostos brancos eram avaliados significativamente mais positivamente do que rostos negros, diferença que foi eliminada após o treino.

Em um estudo subsequente, Mizael et al. (2021) avaliaram três parâmetros específicos derivados da literatura de equivalência (treino misto, redução de feedback e treino explícito de simetria), aplicando cada um a um grupo distinto, além de um grupo controle, para o qual nenhum desses parâmetros foi aplicado. Embora todos os grupos tenham recebido o mesmo treino AB–BC, os parâmetros produziram efeitos distintos: 33 das 46 crianças formaram classes de equivalência, com o grupo Simetria apresentando o maior rendimento e o grupo Controle o menor. As reduções de viés ocorreram quase exclusivamente entre as crianças que formaram as classes, fornecendo forte evidência do vínculo funcional entre responder relacional derivado e mudança avaliativa. A manutenção após seis semanas foi moderada, mas robusta entre aquelas que formaram classes, com muitas preservando pelo menos parte das relações treinadas e derivadas.

Por fim, Passarelli et al. (2024) utilizaram a mesma estrutura AB–BC em um delineamento de medidas repetidas envolvendo sete avaliações por criança: três antes do treino, três após o treino e uma avaliação de seguimento após duas semanas. A generalização foi testada por meio de um teste relacional, avaliações explícitas do tipo Likert (valência emocional), o FAST (uma medida implícita) e o Teste da Boneca. A intervenção produziu mudanças claras nas funções avaliativas em múltiplas medidas, embora a correspondência entre instrumentos tenha variado, especialmente entre avaliações baseadas em seleção de respostas e tarefas de generalização, o que dificulta a interpretação dos efeitos de generalização. Apesar dessas diferenças, o estudo replicou o achado central dos

dois experimentos anteriores: quando classes de equivalência que vinculam rostos negros a símbolos positivos são estabelecidas com sucesso, as crianças apresentam reduções consistentes em avaliações raciais negativas.

#### *Enquadramento Relacional e Reforçamento Diferencial em Adultos*

Dois estudos com adultos utilizaram procedimentos laboratoriais baseados em controle contextual para examinar como pistas relacionais influenciam o responder a estímulos relevantes para o viés racial e microagressões. Em Suárez et al. (2024), os participantes foram selecionados por meio de uma triagem inicial na qual avaliavam sua resposta empática a cenários envolvendo indivíduos de diferentes grupos raciais, associados a eventos positivos ou negativos relacionados aos próprios valores dos participantes; foram incluídos apenas aqueles cujo escore de empatia para algum grupo racial foi pelo menos quatro pontos abaixo do escore atribuído ao próprio grupo racial. O Experimento 1 incluiu dois adultos e o Experimento 2, três. Em ambos, os participantes leram cenários breves envolvendo membros do exogrupo-alvo e avaliaram sua empatia em uma escala de nove pontos. A manipulação crítica envolveu pistas relacionais que enquadravam o exogrupo em moldura de coordenação ou de distinção em relação aos valores pessoais dos participantes. Pistas de coordenação aumentaram de forma consistente as respostas empáticas, enquanto pistas de distinção as reduziram ou produziram padrões contraempáticos, demonstrando que o enquadramento relacional pode alterar sistematicamente o responder emocional direcionado a exogrupos raciais. Uma limitação a considerar é que os procedimentos não incluíram verificação direta de que os participantes respondiam aos alvos como membros do exogrupo, deixando em aberto em que medida a

categorização racial funcionou como variável de controle efetiva do responder empático.

Romani et al. (2024) conduziram um estudo laboratorial de treino por reforçamento diferencial para ensinar profissionais de saúde a discriminar entre exemplos e não exemplos de microagressões raciais. Vinte e seis funcionários de um departamento de psiquiatria hospitalar nos Estados Unidos completaram três levantamentos sequenciais no Qualtrics: uma avaliação de linha de base e duas fases de treino. Cada levantamento continha dois conjuntos de estímulos, cada um com 10 exemplos de microagressões e 10 não exemplos, derivados de Sue et al. (2007). Muitos itens de microagressão tinham como alvo explícito pessoas negras, como supor perigo ao encontrar um homem negro na rua ou presumir habilidades atléticas ao falar com um estudante negro. Esses itens forneceram exemplos claramente definidos, em termos comportamentais, de conteúdo racial antinegro. O delineamento multielemento contrabalanceou qual conjunto cada grupo recebeu treino primeiro. Durante o treino, os participantes receberam elogios imediatos para respostas corretas e feedback corretivo para erros no conjunto-alvo, enquanto o outro conjunto não recebeu feedback. A acurácia na linha de base variou de moderada a variável. Após o reforçamento diferencial, a acurácia nos conjuntos treinados aumentou para acima de 97% para quase todos os participantes, enquanto os conjuntos não treinados apresentaram pouca melhora até receberem treino direto. Os efeitos foram replicados entre os grupos de participantes e conjuntos de estímulos, e o acordo interobservadores foi de 100%. Em conjunto, o estudo demonstra que um treino comportamental breve, baseado em feedback, pode aumentar de forma confiável a discriminação acurada

de microagressões raciais, incluindo aquelas que refletem estereótipos e pressupostos antinegro comuns.

### *Intervenções com Adultos Baseadas na Análise Contextual do Comportamento*

Um conjunto de quatro estudos empregou a Terapia de Aceitação e abordagens contextuais-comportamentais relacionadas para reduzir o viés racial ou aliviar o sofrimento associado ao racismo. Lillis e Hayes (2007) conduziram um estudo piloto em contexto de sala de aula comparando uma intervenção de sessão única baseada em ACT com uma palestra educativa entre 32 estudantes universitários matriculados em disciplinas sobre questões raciais. Utilizando uma sequência contrabalanceada de avaliações repetidas ao longo de cinco pontos temporais, os pesquisadores examinaram mudanças na consciência do viés, aceitação, defusão cognitiva e intenções comportamentais pró-sociais. A sessão baseada em ACT produziu aumentos consistentes nas intenções pró-sociais e em processos compatíveis com ACT imediatamente após a intervenção e em um follow-up de uma semana, enquanto a palestra educativa apresentou efeitos mínimos ou inconsistentes.

West et al. (2013) realizaram um estudo experimental randomizado avaliando se um breve exercício de escrita de clarificação de valores poderia atenuar o estresse relacionado ao racismo em adultos negros norte-americanos. Quatorze participantes ouviram uma vinheta imagética descrevendo uma situação ambígua de discriminação racial antes e depois de completar uma tarefa de escrita de clarificação de valores ou uma tarefa controle. Foram administradas medidas de sofrimento subjetivo e afeto, incluindo a Subjective Units of Distress Scale (SUDS) e o Positive and Negative Affect Schedule (PANAS), bem como

instrumentos de triagem como o Schedule of Racist Events (SRE) e o Vividness of Visual Imagery Questionnaire (VVIQ). Os resultados indicaram grandes reduções no sofrimento subjetivo e reduções moderadas no afeto negativo na condição de clarificação de valores, embora a significância estatística não tenha sido alcançada devido ao tamanho reduzido da amostra. Os participantes da condição controle, que refletiram sobre seus valores menos importantes, apresentaram melhorias mais fracas ou inconsistentes. Uma limitação interpretativa relevante é que o exercício de escrita pode ter funcionado não pela clarificação de valores em si, mas como contexto que alterou o responder subsequente por deseabilidade social: ao refletir publicamente sobre valores importantes, os participantes podem ter ajustado suas respostas na direção do que percebiam como esperado pelo aplicador, tornando difícil isolar o efeito da variável independente pretendida.

Banks et al. (2021) implementaram uma intervenção manualizada de ACT composta por seis sessões com 20 mulheres negras, combinando psicoeducação sobre racismo com exercícios de desfusão, aceitação e autocompaixão. Avaliações pré e pós-intervenção, utilizando medidas validadas para populações afro-americanas, incluindo o Internalized Racial Oppression Scale (IROS), Depression Anxiety Stress Scales (DASS-21), Internalized Shame Scale (ISS) e Strong Black Woman Schema Scale (SBWS), mostraram reduções significativas na opressão racial internalizada, na vergonha internalizada e no sofrimento psicológico. O programa mostrou-se viável e bem aceito, com alta taxa de retenção e melhorias claras nos processos-alvo.

Williams et al. (2020) avaliaram o Racial Harmony Workshop (RHW), uma intervenção comportamental contextual que integrou elementos experienciais da

Functional Analytic Psychotherapy (FAP) e da ACT com o objetivo de reduzir microagressões e promover conectividade interracial entre estudantes universitários negros e brancos. Quarenta e quatro participantes foram alocados para o RHW ou para uma condição controle ativa envolvendo um documentário e diálogo estruturado. Em múltiplas medidas, incluindo afeto, allophilia, ansiedade intergrupar, atitudes raciais e cognições microagressivas, o RHW produziu ganhos mais robustos entre participantes brancos, que apresentaram aumento de afeto positivo, melhora nas atitudes em relação a pessoas negras e reduções em pensamentos e intenções comportamentais microagressivas. Ambas as condições reduziram o racismo simbólico e o racismo daltônico (*colorblind racism*), e os participantes negros em ambos os grupos apresentaram aumento na identidade racial.

### **Discussão e Direções Futuras**

Esta revisão revela que, apesar do crescente interesse conceitual no racismo antinegro, as investigações empíricas no âmbito da análise do comportamento permanecem limitadas tanto em escopo quanto em continuidade. Ao longo de quase duas décadas, apenas vinte estudos empíricos foram identificados, iniciando-se com o estudo piloto baseado em ACT de Lillis e Hayes (2007). Embora esse trabalho tenha representado um ponto de partida importante, ele não foi seguido por replicações diretas ou extensões programáticas, e a área entrou em um período prolongado sem pesquisas empíricas voltadas ao viés racial. Essa descontinuidade reflete um padrão mais amplo já apontado por estudiosos da análise do comportamento: embora o campo historicamente manifeste preocupação ética com questões sociais, incluindo racismo e

desigualdade, frequentemente encontra dificuldades em traduzir essas preocupações em programas empíricos sustentados de pesquisa (Matsuda et al., 2020; de Souza Silva et al., 2022). Assim, a análise do comportamento apenas recentemente começou a preencher uma lacuna empírica de décadas, produzindo estudos isolados em vez de uma tradição cumulativa coerente capaz de orientar teoria, mensuração e intervenção.

Paradoxalmente, essa escassez empírica contrasta com um recente aumento de publicações conceituais, incluindo números especiais dedicados ao racismo e à equidade, como a edição de 2022 da *Behavior Analysis in Practice* sobre antirracismo. Embora tais contribuições tenham ampliado a visibilidade do tema e estimulado a autorreflexão, a maioria enfatizou discussões normativas em detrimento da análise experimental, com poucas exceções, como Catagnus et al. (2022). De forma importante, Sidman (1994) alertou que prescrições conceituais não apoiadas por análises funcionais correm o risco de enfraquecer o rigor analítico que caracteriza a ciência do comportamento. Um engajamento comportamental significativo com a justiça racial, portanto, requer tanto compromisso ético quanto uma base empírica robusta.

Entre os estudos empíricos revisados, o conjunto de trabalhos fundamentados na Relational Frame Theory (RFT) representa a contribuição metodológica e conceitual mais consistente para a compreensão do viés racial na análise do comportamento. Estudos baseados no Implicit Relational Assessment Procedure (IRAP) demonstraram de forma consistente que o responder relacional baseado em latência revela padrões de viés racial que não são evidentes em medidas explícitas. Barnes-Holmes et al. (2010), por exemplo, mostraram que o

responder pró-branco e antinegro persistiu mesmo quando os participantes acreditavam que suas respostas eram públicas, indicando que preocupações com avaliação social não suplantaram redes relacionais previamente estabelecidas. O viés tornou-se ainda mais forte sob critérios de latência mais curtos, sugerindo que o responder relacional imediato exerceu maior controle sobre o desempenho do que pistas contextuais associadas à avaliação pública. Drake et al. (2015) demonstraram adicionalmente que o IRAP detecta padrões confiáveis de viés racial e é capaz de diferenciar favoritismo intragrupal entre grupos raciais, um nível de precisão funcional amplamente indisponível em medidas de autorrelato. Power et al. (2017a, 2017b) ampliaram essa linha de pesquisa ao mostrar que o responder relacional implícito varia sistematicamente entre participantes negros e brancos, indicando que o viés racial reflete histórias de aprendizagem específicas da comunidade, e não atributos individuais estáveis.

O Function Acquisition Speed Test (FAST) também emergiu como uma medida comportamental promissora. Embora sua aplicação ao viés racial seja relativamente recente (ver Passarelli, Mizael et al., 2024; Rodrigues et al., 2022; Roche et al., 2025), um corpo crescente de evidências indica que o FAST é altamente sensível ao responder relacional breve e à força das relações verbais moldadas por histórias de aprendizagem (ver Cartwright et al., 2016; Cummins et al., 2019; Lalor & Roche, 2025). Amostras de adultos e crianças demonstraram viés antinegro consistente no desempenho no FAST (Passarelli, Mizael et al., 2024; Roche et al., 2025). De forma importante, administrações repetidas com crianças mostraram estabilidade temporal, sugerindo que o FAST pode indexar repertórios avaliativos duráveis que dificilmente seriam captados por introspecção

ou estratégias de autoapresentação. Em conjunto, essas medidas relacionais implícitas indicam que o viés racial é frequentemente governado por histórias de responder relacional derivado, e não por intenções explícitas.

As pesquisas de intervenção, embora ainda preliminares, apresentam efeitos encorajadores. Estratégias baseadas na Terapia de Aceitação e Compromisso têm sido utilizadas para trabalhar clarificação de valores, desfusão e tomada de perspectiva, reduzindo o sofrimento emocional associado ao racismo (West et al., 2013) e a opressão racial internalizada (Banks et al., 2021), além de modificar intenções comportamentais relacionadas ao preconceito (Lillis & Hayes, 2007; Williams et al., 2020). Abordagens de treino relacional, incluindo procedimentos baseados em equivalência com crianças (de Carvalho & de Rose, 2014; Mizael et al., 2016; Mizael et al., 2020), também demonstraram capacidade de reatribuir funções positivas a faces negras e alterar o responder relacional enviesado. Embora os efeitos tenham variado, especialmente em estudos envolvendo estímulos socialmente carregados, esses achados sustentam a possibilidade de que contingências cuidadosamente arranjadas possam reorganizar redes relacionais pré-experimentais.

Uma lição central derivada da sequência de estudos com treino baseado em equivalência não é que classes socialmente carregadas sejam inerentemente resistentes à mudança, mas sim que a robustez do procedimento de treino é decisiva. No primeiro estudo, o protocolo AB-BC foi insuficiente para enfraquecer as fortes redes raciais pré-experimentais que controlavam o desempenho das crianças, apesar de todas terem dominado as relações treinadas (de Carvalho & de Rose, 2014). Nos três experimentos subsequentes, entretanto, refinamentos

metodológicos, como sequências simples-para-complexas e treino misto, produziram de forma consistente a emergência e a manutenção de classes de equivalência que vinculavam faces negras a símbolos positivos. Esses achados convergentes indicam que molduras relacionais pré-experimentais podem ser sobrepostas quando o treino inclui parâmetros reconhecidamente eficazes para fortalecer a formação de classes. Assim, em vez de sugerir que procedimentos baseados em equivalência são incapazes de alterar relações socialmente carregadas, a trajetória desenvolvimental desses estudos demonstra que arranjos relacionais mais robustos são capazes de produzir mudanças sistemáticas no responder enviesado.

Aproximadamente 65% dos estudos identificados nesta revisão foram publicados nos últimos cinco anos, indicando um aumento acentuado do engajamento da análise do comportamento com o tema do racismo. Esse crescimento recente parece refletir, ao menos em parte, a influência de Matsuda et al. (2020), cujo artigo examinou explicitamente se a análise do comportamento pode contribuir para a compreensão e a redução do racismo. Ao enquadrar o racismo como um fenômeno comportamental moldado por contingências culturais, esse trabalho funcionou como um chamado à ação, encorajando pesquisadores a abordar um tema que havia sido historicamente sub-representado na pesquisa analítico-comportamental. A produção empírica resultante tem sido publicada predominantemente em periódicos com fortes tradições relacionais e contextuais, como *The Psychological Record* e o *Journal of Contextual Behavioral Science*, em vez de veículos como *Journal of Applied Behavior Analysis* ou *Journal of Experimental Analysis of Behavior*. Esse padrão provavelmente reflete tanto a

compatibilidade conceitual e metodológica das abordagens relacionais com a pesquisa emergente sobre viés racial quanto o caráter preliminar de muitos dos estudos existentes, que frequentemente utilizam amostras pequenas e delineamentos experimentais piloto. Em conjunto, essas tendências sugerem que o campo está começando a ampliar sua agenda científica ao integrar o racismo à investigação analítico-comportamental contemporânea, ao mesmo tempo em que desenvolve gradualmente a sofisticação metodológica necessária para publicação nos periódicos mais tradicionais.

Uma limitação metodológica recorrente nos estudos revisados é a dependência de amostras WEIRD (Western, Educated, Industrialized, Rich, Democratic), um padrão que não é exclusivo da análise do comportamento, mas amplamente reconhecido na ciência psicológica de modo geral (Henrich et al., 2010). No entanto, na análise do comportamento, a questão da representatividade assume uma nuance particular. Historicamente, o campo foi criticado pelo uso de poucos sujeitos, por vezes um único organismo, o que levou a alegações iniciais de que a pesquisa analítico-comportamental carecia de generalidade. Essas críticas contribuíram para a criação do JEAB à medida que Skinner (1956) e Sidman (1960) argumentaram que a validade da ciência do comportamento não reside na representatividade estatística, mas na replicação ordenada de relações funcionais entre variáveis dependentes e independentes.

Ainda assim, a maioria dos estudos incluídos nesta revisão não empregou delineamentos experimentais tradicionais de sujeito único e, em vez disso, adotou delineamentos de grupo, métodos correlacionais ou levantamentos transversais, abordagens que representam uma forma de "empréstimo metodológico" de outras

tradições de pesquisa (Critchfield, 2011; Mace & Critchfield, 2010). Como essas abordagens se baseiam em comparações estatísticas e dados agregados, elas exigem salvaguardas metodológicas distintas daquelas associadas à lógica de caso único, deslocando o foco para questões como amostragem, equivalência entre grupos e validade das conclusões estatísticas (Kazdin, 2021; Cook et al., 2002). Para estudos que utilizam esses métodos, aspectos como viés de não resposta, evasão amostral, inequivalência basal e variáveis de confusão não controladas tornam-se centrais para a validade dos achados. Mais fundamentalmente, do ponto de vista da análise do comportamento, a variabilidade nos dados não é um ruído a ser mascarado por procedimentos estatísticos, mas um produto de variáveis não controladas que devem ser identificadas e estabilizadas experimentalmente (Sidman, 1960). Quando o fluxo de recrutamento, os critérios de exclusão e os padrões de desistência não são relatados, a interpretabilidade dos dados é comprometida de maneiras que não se aplicariam a experimentos com amostras pequenas ou mesmo com um único participante.

Assim, embora a lógica analítico-comportamental não exija representatividade no sentido estatístico (Sidman, 1960), quando pesquisadores adotam metodologias baseadas em grupos ou correlacionais, assumem também as obrigações metodológicas associadas a essas abordagens (Johnston & Pennypacker, 2009). Reconhecer essa distinção é essencial para avaliar adequadamente as forças e limitações da literatura atual e para orientar pesquisas futuras em direção a maior transparência no recrutamento, melhor documentação

da evasão amostral e replicações mais sistemáticas em diferentes contextos e populações.

Uma limitação conceitual adicional observada nos estudos foi a definição inconsistente de viés racial, com muitos trabalhos utilizando termos como “atitudes”, “preconceito” ou “racismo” sem traduzi-los em processos comportamentais. Embora pesquisas recentes tenham começado a articular definições funcionais fundamentadas em responder relacional e transformação de funções de estímulos (Mizael & de Rose, 2016; Passarelli et al., 2023), esses avanços ainda não foram sistematicamente incorporados aos estudos empíricos.

Em conjunto, os achados desta revisão ressaltam a necessidade de uma agenda analítico-comportamental mais cumulativa, rigorosa e conceitualmente coerente para o estudo do viés racial. Pesquisas futuras devem considerar o uso de amostras mais representativas, protocolos pré-registrados, relato transparente de exclusões e medidas comportamentais robustas que capturem ações funcionalmente relevantes. As intervenções precisam ir além de tarefas laboratoriais breves, avaliando o comportamento em contextos socialmente relevantes nos quais o racismo é efetivamente produzido e experienciado. Se a análise do comportamento pretende contribuir de forma significativa para a equidade racial, será necessário avançar além da discussão conceitual e investir na construção de uma ciência empírica sustentada, capaz de analisar e transformar as contingências que mantêm a desigualdade racial.

-----

## **Comentário**

### **Sobre a Construção de Uma Agenda Sistemática em Pesquisas sobre Racismo e Preconceito Racial na Análise do Comportamento**

Ao concluir esta revisão sistemática, torna-se evidente que o percurso da análise do comportamento no estudo do viés racial antinegro tem sido menos o resultado de um programa empírico planejado e mais uma sequência de esforços pontuais, frequentemente motivados por demandas sociais urgentes e inquietações éticas legítimas. Esse padrão não é exclusivo deste domínio, mas assume aqui uma relevância particular, dada a complexidade conceitual e social do fenômeno em questão.

O conjunto de estudos revisados neste capítulo estabelece que o viés racial pode ser tratado como um fenômeno comportamental analisável, passível de mensuração experimental e modificação sob condições controladas. Ao operacionalizar o preconceito racial em termos de responder relacional e transformação de funções de estímulos, essa literatura demonstra que fenômenos socialmente complexos podem ser incorporados ao escopo analítico-comportamental sem recorrer a explicações mentalistas.

O principal desafio identificado não diz respeito à viabilidade conceitual do fenômeno, mas às contingências que selecionam quais programas de pesquisa se tornam contínuos e cumulativos. Fenômenos como o viés racial impõem custos metodológicos elevados: exigem medidas sensíveis, delineamentos refinados, controle rigoroso de variáveis históricas e, frequentemente, produzem efeitos graduais e heterogêneos. Em contraste, áreas que oferecem maior previsibilidade experimental, aplicação direta e retorno institucional tendem a sustentar programas de pesquisa mais extensos e financeiramente atrativos. Essa assimetria ajuda a explicar por que o estudo do racismo, embora socialmente

relevante, historicamente não se consolidou como um programa empírico lucrativo ou de rápida produção científica.

Ainda assim, os estudos baseados em equivalência de estímulos e em procedimentos relacionais indicam que a dificuldade não é estrutural, mas procedimental. Quando parâmetros experimentais robustos são empregados e quando os estudos são organizados de forma programática, torna-se possível produzir efeitos replicáveis e analisar sistematicamente os limites e alcances da intervenção. O avanço do campo, portanto, depende menos de novas formulações conceituais e mais da continuidade deliberada de programas experimentais, nos quais cada estudo refine o controle estabelecido pelo anterior.

É nesse ponto que se insere o próximo capítulo, que desloca o foco da revisão da literatura para a avaliação empírica de medidas e procedimentos, examinando de forma sistemática a estabilidade, a sensibilidade e a generalidade dos efeitos produzidos. A partir desse movimento, a tese passa a tratar o viés racial não como um problema a ser descrito, mas como um fenômeno a ser experimentalmente analisado.

#####

### **Capítulo 3. Medindo o Preconceito Racial**

#### **Background**

Os estudos conduzidos por Mizael et al. (2016, 2021), em sua dissertação de mestrado e tese de doutorado, representaram um avanço importante ao demonstrar que é possível formar classes de equivalência envolvendo estímulos socialmente e afetivamente carregados, como faces negras e um símbolo positivo. A formação dessas classes esteve relacionada à redução do viés racial e à mitigação das diferenças de valência atribuídas a faces negras e brancas. Além disso, a literatura tem mostrado que determinados parâmetros procedimentais podem facilitar a emergência, a manutenção e a generalização das classes de equivalência.

Esses achados, no entanto, abriram um conjunto de questões que não diziam respeito à eficácia do procedimento em si, mas às suas possibilidades de expansão. Entre elas estavam a generalidade dos efeitos para medidas mais próximas de contextos naturalísticos e a viabilidade de sua translação para atividades típicas da infância, como jogos e narrativas. No entanto, antes de qualquer novo avanço interventivo, tornava-se necessário examinar essas questões por meio de etapas intermediárias de replicação, extensão e refinamento das estratégias de mensuração.

A pesquisa científica, especialmente quando se ocupa de fenômenos socialmente complexos, impõe uma exigência central: evitar saltos conceituais e empíricos. Antes de perguntar como modificar o preconceito racial, era necessário responder a uma pergunta logicamente anterior: como medi-lo de forma consistente dentro de um arcabouço analítico-comportamental? Quais medidas poderiam ser utilizadas? Que tipo de fenômeno cada uma delas capturaria? E, sobretudo, que explicações funcionais poderiam ser oferecidas para o desempenho observado nessas medidas?

As definições apresentadas nos capítulos anteriores indicaram que o alvo de interesse não era simplesmente uma “avaliação negativa”, mas padrões de responder relacional. Isso, por sua vez, colocava novas questões. Estaríamos interessados em avaliações de valência emocional? Em relações de equivalência (coordenação), oposição ou comparação? E como detectar mudanças nesses padrões por meio de medidas que se aproximassem, ainda que parcialmente, de situações cotidianas? Além disso, uma questão metodológica adicional se impunha: a simples exposição repetida a medidas de viés poderia, por si só, produzir efeitos de aprendizagem, comprometendo sua função como instrumentos de mensuração?

No campo da análise experimental do comportamento, a estabilidade das medidas é uma condição básica para qualquer inferência causal: mudanças observadas devem poder ser atribuídas às manipulações experimentais, e não à simples repetição do teste. Foi nesse contexto que surgiu um impasse metodológico concreto.

Quando se estabeleceu a parceria com a primeira escola disposta a participar da pesquisa, havia 34 termos de consentimento livre e esclarecido assinados, 34 termos de assentimento e um grupo de crianças motivadas a participar. Diante desse cenário, duas alternativas estavam disponíveis.

A primeira consistia em seguir rigidamente o protocolo proposto para o projeto de pesquisa: aplicar as medidas que eram critério de inclusão, identificar apenas as crianças que apresentassem viés racial no pré-teste e excluir aquelas que não o apresentassem, iniciando imediatamente o procedimento interventivo. Essa estratégia permitiria um avanço mais rápido em direção ao experimento propriamente dito.

A segunda alternativa implicava suspender temporariamente o início da intervenção para investigar uma questão mais básica: a estabilidade e a sensibilidade das próprias

medidas. Isso incluía examinar se os índices de viés se mantinham ao longo de repetições sucessivas, tanto entre crianças que apresentavam viés na primeira avaliação quanto entre aquelas que não o apresentavam.

A escolha da primeira estratégia teria sido menos custosa em termos de recursos, deslocamento, tempo. No entanto, ela pressupunha que as medidas disponíveis fossem não apenas válidas, mas também estáveis e sensíveis. Diante de uma amostra relativamente ampla e de um conjunto de quatro medidas distintas, optou-se por uma abordagem mais cautelosa. O interesse deslocou-se do simples mapeamento de “quem apresenta viés” para questões mais fundamentais: essas medidas são sensíveis para detectar padrões de viés racial? Elas produzem resultados estáveis ao longo de exposições repetidas? E esse padrão de estabilidade ou instabilidade difere entre indivíduos com e sem viés inicial?

Essas questões orientaram o primeiro estudo empírico desta tese, intitulado *Estabilidade e Sensibilidade de Medidas de Preconceito Racial*, publicado em inglês no *European Journal of Behavior Analysis*. É importante esclarecer, contudo, que nem todos os dados coletados nesse estudo foram reportados no artigo. O foco editorial recaiu sobre um aspecto específico que tem recebido atenção crescente na literatura: a correspondência entre medidas implícitas e explícitas. Por essa razão, o artigo concentrou-se no desempenho no *Function Acquisition Speed Test* (FAST) e em uma medida explícita de valência emocional.

Um aspecto interessante do percurso editorial associado a esse tipo de pesquisa diz respeito às contingências éticas que passam a controlar a avaliação científica quando o objeto de estudo envolve estímulos socialmente sensíveis. Em revisões por pares, observou-se que críticas centrais ao estudo não se concentraram em delineamento experimental, análise de dados ou coerência conceitual, mas na própria legitimidade ética do

procedimento, especificamente no uso de estímulos raciais em pareamentos com imagens de valência negativa.

Ainda que o estudo tenha sido aprovado por comitê de ética, utilizado bases de dados previamente validadas, incluído condições simétricas envolvendo faces brancas e estímulos negativos, e previsto intervenções e procedimentos psicoeducativos subsequentes para todos os participantes, a objeção persistiu. Esse episódio ilustra como, em pesquisas sobre preconceito racial, o escrutínio ético frequentemente ultrapassa critérios procedimentais formais e passa a operar como uma variável adicional que seleciona quais perguntas podem ou não ser investigadas experimentalmente.

Longe de invalidar o estudo, essa reação evidencia uma tensão central do campo: investigar empiricamente o preconceito racial implica, inevitavelmente, expor relações que já operam no ambiente social, ainda que tal exposição produza desconforto ético. Ignorar essas relações não as elimina; apenas desloca seu controle para fora do alcance da análise funcional. Nesse sentido, a controvérsia editorial observada não se restringe a um procedimento específico, mas revela limites sociais e institucionais impostos à própria investigação científica do racismo.

Uma das questões investigadas nesse estudo, e ainda pouco explorada na análise do comportamento, dizia respeito ao efeito da mera exposição às medidas. Em outras palavras, tarefas implícitas como o FAST poderiam gerar aprendizagem suficiente para alterar o desempenho ao longo das repetições, reduzindo sua sensibilidade como instrumentos de mensuração? O mesmo ocorreria com medidas explícitas? Essas perguntas não serão respondidas neste capítulo, pois constituem o foco do estudo subsequente. No entanto, dois conjuntos de dados relevantes, não reportados no artigo publicado, merecem ser

apresentados aqui, pois ajudam a compreender o percurso conceitual e metodológico que se seguiu.

A primeira dessas medidas foi o Teste Relacional, denominado por Mizael et al. (2016, 2020) de Teste de Equivalência Modificado (AC3). O objetivo desse teste era verificar a frequência com que as crianças relacionavam faces negras (C1) e faces brancas (C3) com um símbolo positivo (A1) e um símbolo negativo (A2). O teste recebeu essa denominação porque, diferentemente do teste AC tradicional utilizado no procedimento de intervenção, no qual havia apenas duas comparações (face negra e estímulo abstrato), o AC3 introduzia uma terceira comparação: a face branca.

Essa modificação é conceitualmente relevante. Selecionar uma face negra em associação com um símbolo positivo quando a alternativa é um estímulo abstrato não impõe o mesmo custo relacional que fazê-lo na presença simultânea de uma face branca. Estudos prévios indicam que estímulos mais familiares tendem a ser preferidos, o que torna plausível a formação bem-sucedida de relações positivas com o estímulo negro na ausência de um estímulo racialmente contrastante. A introdução da face branca altera o contexto relacional: a pergunta deixa de ser se o negro pode ser positivo e passa a ser se ele permanece positivo quando o branco está disponível como alternativa numa situação em que só uma opção pode ser escolhida.

Esse teste mostrou-se informativo tanto no pré-teste quanto no pós-teste. Antes da intervenção, ele permitia que a criança evitasse relacionar o negro ao negativo selecionando o estímulo abstrato, mesmo que houvesse uma preferência implícita pelo branco. No pós-teste, a questão tornava-se mais exigente: a relação entre negro-positivo emergiria e se manteria mesmo na presença explícita da face branca?

Os dados indicaram que o teste relacional foi sensível para identificar padrões distintos de responder, incluindo molduras de oposição (por exemplo, se branco é positivo, negro é negativo) e padrões mais flexíveis, nos quais o branco não era selecionado como negativo, mas o negro o era. Além disso, a medida apresentou relativa estabilidade entre crianças que exibiam viés racial consistente: ao longo de três exposições distribuídas em uma semana, não houve mudança significativa no índice médio de viés desse grupo. Em contraste, um subconjunto expressivo de crianças apresentou oscilações consideráveis ao longo das repetições. Em um dia, o negro era selecionado como negativo; em outro, o branco. Esses padrões sugerem um responder relacional altamente dependente de eventos contextuais recentes, indicando maior flexibilidade relacional, ainda que sem um viés estável definido.

A segunda medida não reportada no artigo foi o clássico Teste da Boneca (Clark & Clark, 1947). Embora frequentemente tratado como um instrumento histórico da psicologia social, o teste pode ser descrito, em termos funcionais, como um procedimento relacional: uma pergunta avaliativa verbal funciona como estímulo modelo, e as bonecas funcionam como estímulos de comparação. No presente estudo, duas bonecas idênticas, diferenciadas apenas pela cor da pele, foram apresentadas às crianças, que responderam a cinco perguntas avaliativas positivas (por exemplo, “Quem é boa?”, “Quem é inteligente?”, “Quem é confiável?”) e a cinco perguntas negativas com adjetivos opostos.

Essa medida mostrou-se significativamente menos estável. Mesmo entre crianças que apresentavam viés racial consistente no teste relacional AC3, o Teste da Boneca frequentemente indicou avaliações positivas dirigidas às bonecas negras e negativas às brancas. Esse padrão sugere baixa correspondência entre as medidas e indica possível

influência de contingências sociais de desejabilidade, especialmente por se tratar de uma tarefa explícita, verbal e facilmente interpretável em termos normativos.

Em síntese, o conjunto de medidas revelou um quadro assimétrico. A medida relacional mostrou-se relativamente estável para crianças com viés consistente, enquanto a medida mais próxima de um contexto naturalístico, o Teste da Boneca, apresentou elevada variabilidade. Esses resultados não indicam um fracasso metodológico, mas apontam problemas conceituais e experimentais que precisariam ser enfrentados antes de qualquer avanço interventivo. Essas questões serão retomadas no capítulo seguinte. No presente capítulo, o foco recairá exclusivamente sobre o desempenho no FAST e na escala Likert, preparando o terreno para as análises subsequentes.

\*\*\*

**Estudo 2 - Examinando o viés racial antinegro implícito e explícito em crianças: um delineamento de medidas repetidas utilizando o *Functional Acquisition Speed Test (FAST)*<sup>5</sup>**

Medidas explícitas e implícitas tem sido utilizada para detectar o viés antinegro e avaliar a eficácia de intervenções. A análise do comportamento carece de estudos que utilizem medidas implícitas repetidas para avaliar intervenções sobre viés racial. Assim, este estudo teve como objetivo identificar viés racial antinegro implícito e explícito em uma amostra de crianças e examinar a confiabilidade teste-reteste dessas medidas ao longo de três momentos de avaliação repetidos. Trinta e quatro crianças (M = 8,60 anos) avaliaram imagens de rostos negros e brancos utilizando uma escala Likert de cinco pontos. O *Function Acquisition Speed Test (FAST)*, uma medida implícita, avaliou a fluência com que os participantes categorizavam rostos negros e brancos com imagens positivas e negativas. Os resultados indicaram ausência de viés racial explícito na amostra examinada. No entanto, os resultados referentes ao viés implícito revelaram uma tendência de viés antinegro na amostra geral, com significância observada em 11 participantes que apresentaram consistentemente viés antinegro ao longo das três repetições. Quanto à confiabilidade teste-reteste das medidas, não foram observadas diferenças significativas de desempenho nas medidas implícitas e explícitas entre as três aplicações. Esses achados sugerem que o FAST pode ser uma ferramenta sensível para identificar viés racial em crianças, podendo ser usado como medida repetida.

---

<sup>5</sup> Passarelli, D. A., Mizael, T. M., Silvestre, M. H., & de Rose, J. C. (2024). Examining implicit and explicit anti-Black racial bias in children: A repeated measure design using the Function Acquisition Speed Test (FAST). *European Journal of Behavior Analysis*, 25(2), 153-173.  
<https://doi.org/10.1080/15021149.2024.2388487>

## Introdução

O viés racial é caracterizado por atitudes negativas condicionadas, presentes em uma cultura, em relação a um grupo étnico (Mizael & de Rose, 2017; Nogueira, 2007; Schneider, 2005). Especificamente, o viés antinegro refere-se a atitudes negativas direcionadas a indivíduos de ascendência africana (Corrington et al., 2023). Na literatura da análise do comportamento, atitudes têm sido interpretadas de diferentes formas, incluindo como comportamento verbal (ver Guerin, 1994), relações derivadas (e.g., Grey & Barnes, 1996; Kohlenberg et al., 1991; Mizael et al., 2016) e molduras relacionais (e.g., Hayes et al., 2001). O aspecto avaliativo das atitudes é compatível com todas essas interpretações; portanto, atitudes podem ser compreendidas como avaliações verbais dirigidas a um estímulo relevante (Gawronski, 2007; Hayes et al., 2001).

As atitudes podem ser explícitas ou implícitas, e alguns estudos têm empregado medidas avaliativas para identificar o viés racial e avaliar a eficácia de intervenções (ver Gonzalez et al., 2017; Mizael et al., 2016, 2020; Sinclair et al., 2005; ver também Lai et al., 2014, 2016). No contexto da Análise do Comportamento, avaliações “explícitas” são conceitualizadas como respostas relacionais estendidas e elaboradas, emitidas sem restrição temporal (Barnes-Holmes et al., 2010). Por exemplo, expressar abertamente afirmações como “não gosto de pessoas negras porque são preguiçosas” representa uma avaliação explícita. Métodos para avaliar atitudes explícitas incluem questionários (e.g., Santos et al., 2006), escalas de avaliação do tipo Likert (e.g., Bradley & Lang, 1994; Osgood et al., 1957) e testes relacionais (e.g., Clark & Clark, 1947; Mizael et al., 2016, 2020).

Por outro lado, avaliações “implícitas” são caracterizadas por respostas relacionais breves e imediatas, emitidas dentro de um intervalo de tempo restrito (Barnes-Holmes et al., 2010). Por exemplo, desviar-se rapidamente ao perceber a presença de uma pessoa negra pode ser considerado uma avaliação implícita. As atitudes implícitas são medidas por procedimentos que acessam a história de respostas relacionais, geralmente impondo restrição temporal às respostas.

Diferentes instrumentos podem ser utilizados para avaliar atitudes implícitas. Na psicologia social, o Implicit Association Test (IAT; Greenwald, 1999) é um dos mais utilizados. Na análise do comportamento, o Implicit Relational Assessment Procedure (IRAP; Barnes-Holmes et al., 2010) e o Functional Acquisition Speed Test (FAST; O'Reilly et al., 2012) estão entre os procedimentos mais empregados. Nesses instrumentos, a noção de medida implícita não se refere à topografia da resposta em si, que é observável e operante em ambos os casos, mas a propriedades da resposta: no IRAP, a latência é a medida central; no FAST, embora a medida seja a velocidade de aquisição de relações condicionais expressa na inclinação da curva acumulada, latências menores contribuem diretamente para inclinações mais acentuadas. O termo “implícito” designa, portanto, que essas propriedades do responder escapam ao controle verbal direto do participante, não que a resposta em si seja encoberta.

O FAST foi desenvolvido para avaliar a fluência de relações aprendidas pelos indivíduos. O instrumento é tipicamente composto por quatro blocos de discriminações simples, incluindo dois blocos de linha de base com relações arbitrárias, um bloco consistente com estereótipos assumidos e um bloco inconsistente com esses estereótipos. Esses blocos são apresentados com

instruções mínimas, e os participantes recebem feedback tanto para respostas corretas quanto incorretas em cada tentativa, dependendo do bloco. Esse feedback pode facilitar a aprendizagem e reforçar respostas acuradas.

Durante o procedimento, os estímulos são apresentados sequencialmente, um de cada vez, na tela do computador, e os participantes respondem de forma consistente ou inconsistente com base em sua história de aprendizagem. Por exemplo, no contexto da avaliação de viés racial, os participantes recebem feedback de “correto” ao pressionar a tecla “z” na presença de rostos negros e também na presença de palavras negativas. De modo semelhante, recebem feedback de “correto” ao pressionar a tecla “m” diante de rostos brancos e palavras positivas. Respostas incorretas, como pressionar a tecla “z” na presença de rostos brancos ou exceder o limite de tempo de 3000 milissegundos, resultam em feedback de “incorreto”. No bloco inconsistente, essa contingência é invertida, de modo que “z” para rostos brancos e palavras negativas, e “m” para rostos negros e palavras positivas, passam a receber feedback de “correto”. Ao comparar a fluência com que os indivíduos adquirem padrões de resposta consistentes e inconsistentes com sua história de aprendizagem, o FAST fornece uma medida da taxa de aquisição de relações entre esses estímulos.

O formato do FAST passou por refinamentos ao longo de uma série de estudos empíricos progressivos (Cartwright et al., 2016; Cummins & Roche, 2020; Cummins et al., 2018, 2019; O’Reilly et al., 2013), posicionando-o como uma alternativa para analistas do comportamento interessados em avaliar a velocidade de aquisição de classes funcionais de respostas. O FAST se destaca por sua aplicação simples e rápida, o que minimiza perdas amostrais — uma característica

também observada em outros instrumentos (ver Geist et al., 2023; Errasti et al., 2019). Tipicamente, uma aplicação do FAST dura cerca de 10 minutos, o que é particularmente vantajoso ao trabalhar com crianças pequenas, que podem ter dificuldade em manter o engajamento em tarefas que exigem atenção prolongada (Danis et al., 2008; Gautier & Droit-Volet, 2002).

Além disso, sua versatilidade permite sua inclusão como medida avaliativa em procedimentos que envolvem intervenções, demonstrando sua adaptabilidade a diferentes contextos experimentais. Por meio do cálculo da inclinação da taxa de aquisição de classes funcionais de resposta para cada bloco do FAST, o método acompanha sistematicamente, tentativa a tentativa, os aumentos de fluência de resposta em tempo real, utilizando um procedimento semelhante a registros cumulativos, sem excluir parâmetros temporais do procedimento do cálculo (ver Cummins et al., 2018 para uma discussão sobre métodos alternativos de pontuação).

A eficácia do FAST em identificar relações verbais avaliativas em diferentes contextos já foi demonstrada. Pesquisas sociais anteriores utilizaram o FAST para investigar taxas de aprendizagem em relações entre mulheres e estereótipos de gênero (Cartwright et al., 2016) e entre tons de pele mais escuros e adjetivos negativos (Rodrigues et al., 2022). Por exemplo, o estudo de Cartwright et al. (2016) investigou estereótipos de gênero tipicamente associados à masculinidade (e.g., competitividade e dominância) e à feminilidade (e.g., cuidado e delicadeza). Os autores utilizaram uma versão modificada do FAST, na qual os blocos de linha de base foram removidos, sendo apresentados um bloco de prática e dois blocos de teste (um consistente e outro inconsistente). No bloco de prática, estímulos não

relacionados ao estudo foram apresentados para familiarizar os participantes com o procedimento. O bloco consistente envolveu o fornecimento da mesma resposta para relações funcionais compatíveis com estereótipos tradicionais (e.g., pressionar a mesma tecla para estímulos femininos e estereótipos de feminilidade). Em contraste, o bloco inconsistente exigiu respostas comuns para estímulos masculinos e estereótipos de feminilidade, e para estímulos femininos e estereótipos de masculinidade. Os participantes também responderam a medidas de autorrelato de viés de gênero. Os resultados indicaram que todos os participantes apresentaram taxas de aprendizagem mais rápidas no bloco consistente com estereótipos de gênero. Além disso, os autores observaram ausência de correlação entre as medidas implícitas e explícitas de autorrelato (para resultados semelhantes, ver Nosek, 2007; Barnes-Holmes et al., 2010; O'Reilly et al., 2015).

No contexto do viés racial, um estudo conduzido por Rodrigues et al. (2022) investigou a aplicação do FAST modificado (semelhante ao de Cartwright et al., 2016) para avaliar histórias de aprendizagem verbal relacionadas a estímulos raciais em uma amostra de 81 estudantes universitários brasileiros. Os estímulos experimentais consistiram em imagens de rostos femininos digitalmente alterados para representar diferentes tons de pele (mais claros e mais escuros) e palavras com valência positiva ou negativa. O estudo investigou se a variável “tom de pele”, isolada de outras características étnicas, poderia funcionar como estímulo discriminativo para atitudes de viés antinegro, mesmo quando os modelos não eram originalmente negros. Além disso, buscou-se avaliar a utilidade do FAST para mensurar a história de aprendizagem verbal em temas envolvendo viés racial

direcionado a tons de pele mais escuros. Os resultados indicaram que 63% dos participantes (51 de 81) apresentaram taxas de aprendizagem mais rápidas no bloco consistente com tons de pele mais escuros e palavras de valência negativa. Esses achados sugerem que o FAST apresenta sensibilidade para quantificar o viés racial utilizando estímulos raciais isolados.

Apesar da noção predominante de que o preconceito racial e a discriminação estão majoritariamente presentes em adultos, pesquisas no campo do desenvolvimento humano indicam que crianças a partir de quatro anos de idade já podem apresentar viés racial antinegro (Aboud, 1988; Bar-Tal, 1966; Clark & Clark, 1947). Compreender o preconceito racial implícito em crianças é fundamental por diversas razões. Primeiramente, as interações verbais e os ambientes aos quais as crianças são expostas desde cedo (ver Aboud, 1993; Auty & Lewis, 2004; Skinner, 1957) podem influenciar significativamente o desenvolvimento de suas atitudes em relação a diferentes grupos étnicos (Blackson et al., 2022; Žeželj et al., 2015). Ao investigar preconceitos implícitos ou explícitos em crianças, é possível compreender como essas atitudes se formam e emergem precocemente ao longo do desenvolvimento.

Além disso, a investigação dessas atitudes permite avaliar em que medida relações verbais já estão estabelecidas na infância, possibilitando o desenvolvimento de estratégias educacionais e programas de intervenção voltados à redução dessas atitudes negativas enquanto as crianças ainda são jovens. Nesse sentido, a investigação de medidas implícitas e explícitas é importante não apenas para identificar o preconceito racial, mas também para avaliar a eficácia de intervenções. Ademais, abordar vieses raciais em crianças é

crucial para o próprio bem-estar infantil, uma vez que atitudes negativas podem levar a dificuldades sociais, emocionais e acadêmicas, afetando o desenvolvimento global e as relações interpessoais (Bigler & Liben, 2007; Cvencek et al., 2016). Programas de intervenção e educação precoce voltados ao viés racial podem contribuir para mitigar esses efeitos negativos e promover uma sociedade mais inclusiva e equitativa.

Numerosos estudos de diferentes áreas da psicologia investigaram tanto o viés racial antinegro explícito quanto implícito em crianças (Carvalho & de Rose, 2014; Gonzalez et al., 2017; Katz et al., 1975; Mizael et al., 2016, 2020; Nesdale, 1999; Nesdale et al., 2005; Setoh et al., 2019; Sinclair et al., 2005; ver Slobodin et al., 2023 para uma revisão sistemática). Na literatura da análise do comportamento, um número limitado de estudos examinou o viés racial implícito e explícito (Barnes-Holmes et al., 2010; Power et al., 2010; Rodrigues et al., 2022; ver também Farmer et al., 2023 para uma revisão sistemática). Notavelmente, há uma lacuna de pesquisas que empreguem essas medidas em crianças para investigar o viés racial (ver, contudo, Mizael et al., 2016).

Uma exceção relevante é o estudo de Mizael e colaboradores (2016). Os autores selecionaram 13 crianças que consistentemente selecionavam rostos negros na presença de símbolos negativos e atribuíram escores mais baixos a rostos negros em comparação a rostos brancos em uma escala Likert. Os autores implementaram uma intervenção baseada no paradigma de equivalência de estímulos, treinando os participantes a relacionar indiretamente rostos negros a estímulos positivos. Após a intervenção, os participantes deixaram completamente de relacionar rostos negros a símbolos negativos. Além disso, a diferença

significativa observada antes do procedimento na avaliação explícita de rostos negros e brancos tornou-se não significativa após a intervenção. Adicionalmente, por meio do *Implicit Relational Assessment Procedure* (IRAP), uma medida implícita de atitudes, os autores avaliaram propriedades avaliativas em blocos consistentes e inconsistentes com o viés antinegro. Os resultados indicaram ausência de viés racial significativo nas crianças; no entanto, a ausência de uma medida IRAP no pré-teste dificulta a determinação de se esse resultado ocorreu em função da intervenção implementada (Mizael et al., 2020).

O presente estudo teve como objetivo avaliar a eficácia do FAST na identificação do preconceito racial em uma amostra de crianças e explorar a correlação entre medidas implícitas (FAST) e explícitas (escala do tipo Likert). A escolha do FAST para esta investigação decorre da aplicação mais rápida do instrumento e de seus procedimentos de discriminação simples, o que o torna adequado para essa faixa etária e minimiza potenciais perdas amostrais (ver Watters et al., 2023). Outro aspecto relevante do presente estudo refere-se à preocupação recorrente com a confiabilidade teste-reteste de medidas implícitas quando aplicadas repetidamente. Há uma lacuna de conhecimento sobre a dinâmica dessas medidas ao longo de aplicações repetidas no campo da análise do comportamento.

Como mencionado anteriormente, Mizael et al. (2016) empregaram uma medida implícita após uma intervenção comportamental voltada à mitigação do viés racial. Contudo, a ausência dessa medida no pré-teste impede uma determinação conclusiva dos efeitos da intervenção sobre o viés racial antinegro implícito. Assim, como objetivo exploratório secundário, este estudo busca

investigar mudanças nas medidas implícitas e explícitas ao longo de aplicações repetidas, examinando especificamente a possível influência do efeito teste-reteste no desempenho dos participantes ao longo de três aplicações espaçadas.

## Método

### Participantes e Local

Trinta e quatro participantes (21 do sexo feminino; *idade média* = 8,6 anos), provenientes de uma escola municipal pública no estado de São Paulo, Brasil, foram recrutados para o estudo. Foi empregado um delineamento de medidas repetidas intrassujeitos. A análise de poder estatístico, conduzida por meio do software G\*Power (Faul et al., 2007), indicou um tamanho de efeito médio ( $f^2 = 0,26$ ), com nível de significância de 0,05 e poder estatístico de 95%. A Tabela 1 apresenta a caracterização da amostra quanto ao gênero e à raça dos participantes.

Os responsáveis legais relataram a raça das crianças com base nas classificações adotadas pelo Instituto Brasileiro de Geografia e Estatística (IBGE). É importante destacar que, no Brasil, indivíduos com ascendência racial mista, como aqueles com ancestralidade negra e branca, são comumente classificados como “pardo”. Os dados foram coletados na biblioteca de uma escola municipal localizada em uma cidade de pequeno porte no estado de São Paulo, Brasil.

### Tabela1.

#### *Caracterização da Amostra*

Gênero	Raça	N (%)
Feminino	Branco	11 (32.35%)
	Preto	4 (11.76%)
	Pardo	6 (17.64%)
Masculino	Branco	2 (5.8%)
	Preto	3 (8.82%)
	Pardo	8 (23.53%)

## Procedimento

O procedimento foi aprovado pelo Comitê de Ética em Pesquisa da Universidade Federal de São Carlos (CAAE nº 48028721.0.0000.5504) e pela Secretária Estadual de Educação do Estado de São Paulo. A participação foi voluntária, com consentimento dos responsáveis legais e assentimento das crianças, sendo os estudos conduzidos mediante autorização da Secretaria Pública de Educação e da direção das escolas envolvidas.

Foi conduzido um delineamento de medidas repetidas, no qual duas medidas de viés racial foram obtidas em três momentos distintos ao longo de uma semana. As medidas foram aplicadas com intervalo de dois dias, nas segundas-feiras, quartas-feiras e sextas-feiras. A coleta de dados foi realizada por uma pesquisadora, mulher branca, de 30 anos, em um espaço silencioso e controlado da biblioteca da escola, com o objetivo de minimizar possíveis influências externas. Para garantir padronização e evitar distrações, apenas um participante era chamado por vez para a realização do procedimento.

Todo o procedimento teve duração entre 12 e 15 minutos para cada criança, e as medidas foram coletadas na seguinte ordem:

**Escala de Valência Emocional** (Bradley & Lang, 1994 – adaptada do *Self-Assessment Manikin*). Para a avaliação explícita de rostos negros e brancos, foi utilizada uma escala tipo Likert de cinco pontos, com expressões emocionais representadas por emojis. A escala foi programada no software PsychoPy 3.0, de modo que a imagem de um rosto negro ou branco aparecia no centro da tela, com a escala posicionada abaixo da imagem. Os participantes receberam instruções verbais para fornecer uma avaliação positiva, neutra ou negativa da imagem, com

base nas emoções que o rosto evocava. Avaliações entre 1 e 2 indicaram valência negativa, o ponto 3 foi considerado neutro, e avaliações entre 4 e 5 indicaram valência positiva.

A versão do ***Function Acquisition Speed Test*** (FAST – Cartwright et al., 2016) utilizada foi desenvolvida por meio do software de programação LiveCode™ e administrada em um computador portátil com tela de 16 polegadas (resolução de 1920 × 1080 pixels, taxa de atualização de 120 Hz). O protocolo FAST empregado neste experimento diferiu parcialmente de estudos previamente publicados (e.g., O'Reilly et al., 2012, 2013), por consistir em apenas um bloco de prática e dois blocos de teste, com a omissão dos blocos de linha de base.

Antes do início da tarefa, a experimentadora explicou verbalmente a natureza geral do procedimento, utilizando as seguintes instruções padronizadas do FAST (ver Cartwright et al., 2016; O'Reilly et al., 2012):

“Na próxima etapa, sua tarefa será aprender qual botão pressionar quando uma imagem aparecer na tela. **IMPORTANTE:** durante esta fase, você deve pressionar apenas a tecla ‘Z’ ou a tecla ‘M’. Por favor, localize essas teclas no teclado agora. Para ajudá-lo(a) a aprender, você receberá feedback informando se sua resposta está correta ou incorreta. Se tiver alguma dúvida, pergunte agora à pesquisadora. Pressione qualquer tecla quando estiver pronto(a) para começar.”

Após o participante pressionar uma tecla aleatória, foi apresentado um bloco de prática composto por 16 tentativas, seguido de dois blocos de teste, cada um contendo 50 tentativas. No bloco de prática, foram apresentadas imagens neutras (300 × 400 pixels) de categorias de estímulos não relacionadas ao estudo

(e.g., partes do corpo, roupas), com o objetivo de familiarizar os participantes com a tarefa sem introduzir vieses.

Nos blocos de teste, foram utilizados estímulos-alvo compostos por quatro rostos negros e quatro rostos brancos, com distribuição equilibrada entre rostos masculinos e femininos. Esses estímulos foram obtidos de Mizael e colaboradores (2016, 2020) e estão disponíveis em <http://faceresearch.org/>. Como atributos avaliativos, foram selecionadas quatro imagens com valência emocional negativa (e.g., um cachorro bravo, um vaso sanitário sujo, uma cobra atacando, uma criança no dentista com expressão de dor) e quatro imagens com valência emocional positiva (e.g., filhotes de cachorro, golfinhos, um bebê e um gatinho). As imagens com valência positiva e negativa foram extraídas do *International Affective Picture System* (IAPS – Lang et al., 2005), um conjunto padronizado amplamente utilizado em pesquisas psicológicas para evocar respostas de valência, excitação e dominância (Bradley & Lang, 1994).

No total, 16 estímulos (oito rostos e oito imagens avaliativas) foram apresentados, cada um aparecendo uma vez por tentativa, com três ou quatro apresentações totais por estímulo ao longo dos blocos. A sequência de apresentação foi randomizada pelo computador.

Os blocos de teste, consistente e inconsistente<sup>6</sup>, foram idênticos em formato, diferindo apenas nas contingências de reforçamento (feedback corretivo)

---

<sup>6</sup>A denominação dos blocos como “consistente” e “inconsistente” refere-se a um estereótipo presumidamente prevalente na sociedade. Considerando que o viés antinegro permanece difundido no Brasil, conforme indicado pela literatura (Almeida, 2019; de Souza et al., 2022), o bloco no qual os participantes eram reforçados por categorizar estímulos negros como negativos e estímulos brancos como positivos foi denominado “consistente”. Em contraste, o bloco no qual os participantes eram reforçados por categorizar estímulos negros como positivos e estímulos brancos

aplicadas. O feedback corretivo (a palavra “correto” ou “incorreto”) era apresentado no centro da tela por 1,5 segundo, em fonte vermelha, tamanho 48. No bloco consistente, os participantes recebiam feedback de “correto” ao pressionar a tecla “Z” na presença de rostos negros e imagens negativas, e a tecla “M” na presença de rostos brancos e imagens positivas, configuração compatível com respostas relacionais indicativas de viés antinegro. Respostas incorretas (e.g., pressionar “M” para rostos negros no bloco consistente) ou respostas emitidas após o limite de tempo de três segundos resultavam em feedback de “incorreto”.

No bloco inconsistente, essa relação era invertida: a tecla “Z” era considerada correta para rostos negros e imagens positivas, e a tecla “M” para rostos brancos e imagens negativas, configuração inconsistente com o estereótipo presumido. Cada estímulo era seguido por um intervalo entre tentativas de 250 ms. A ordem de apresentação dos blocos consistente e inconsistente foi randomizada entre participantes para controlar possíveis efeitos de ordem. Ao final da tarefa, os participantes receberam uma breve mensagem informando o término da atividade e agradecendo sua participação.

O desempenho no FAST foi analisado com base na inclinação do registro cumulativo de respostas corretas em função do tempo de resposta. Essa medida permite avaliar a rapidez com que os participantes aprendem a relacionar corretamente os estímulos de acordo com as regras de cada bloco. Uma inclinação mais acentuada indica maior fluência, refletindo maior número de

---

como negativos foi denominado “inconsistente”. É importante destacar que todos os participantes foram expostos tanto ao bloco consistente quanto ao bloco inconsistente ao longo do teste.

respostas corretas em menor intervalo de tempo. O efeito FAST foi calculado subtraindo-se a inclinação da curva de aprendizagem do bloco inconsistente da inclinação do bloco consistente. A configuração que resulta em maior taxa de aprendizagem indica o padrão de respostas relacionais mais alinhado à história verbal social do participante. Nesse contexto, valores positivos indicam predominância do bloco consistente, sugerindo maior fluência em relações funcionais compatíveis com viés antinegro. Por outro lado, valores negativos indicam maior fluência no bloco inconsistente, refletindo relações funcionais compatíveis com viés pró-negro.

### **Análise de Dados**

Foram conduzidas análises descritivas e inferenciais no software RStudio (versão 2023.06.2+561) para Windows, utilizando os pacotes *datarium*, *rstatix*, *dplyr*, *magrittr* e *multcomp*. As análises descritivas incluíram médias, desvios-padrão e diferenças entre avaliações de rostos brancos e negros, bem como as diferenças de inclinação entre os blocos Consistente e Inconsistente (Efeito FAST) em cada momento de repetição.

De forma relevante, as diferenças individuais dos participantes foram utilizadas para estratificar os dados em subgrupos que apresentaram estabilidade em viés pró-negro ou antinegro, bem como em subgrupos que exibiram alternância desses padrões ao longo do tempo.

Os dados de todos os participantes foram ajustados a um modelo linear por meio da função *lm*. Em seguida, foi realizada uma ANOVA de medidas repetidas com 1000 permutações (Tipo III), procedimento sensível a violações de normalidade e adequado para amostras pequenas. O *eta* quadrado foi calculado

para cada fator, com o objetivo de avaliar a magnitude e a relevância prática das diferenças observadas. Testes *post hoc* com correção de Bonferroni foram conduzidos para comparar os níveis dos fatores em relação à variável dependente, sendo essas análises repetidas para os subgrupos que apresentaram viés estável ou alternância ao longo dos diferentes momentos de avaliação.

## **Resultados**

### **Escala de Valência Emocional**

De modo geral, os rostos negros e brancos receberam avaliações semelhantes, com os rostos negros sendo avaliados de forma ligeiramente mais positiva do que os rostos brancos (ver Tabela 1). A ANOVA de medidas repetidas do Tipo III (Valência: raça do rosto × momento) não indicou diferenças significativas na valência atribuída a rostos negros e brancos,  $F(1) = 0,29$ ,  $p = 0,85$ ,  $\eta^2 = 0,001$ . Da mesma forma, não foram observadas diferenças significativas ao se considerar apenas os três momentos de aplicação da medida,  $F(2) = 0,21$ ,  $p = 0,81$ ,  $\eta^2 = 0,006$ .

Além disso, não houve efeito de interação significativo entre a raça do rosto (negro vs. branco) e os três momentos de avaliação,  $F(2) = 0,01$ ,  $p = 0,98$ ,  $\eta^2 = 0,000$ . Esses resultados indicam que as diferenças entre as médias das avaliações de rostos brancos e negros não foram estatisticamente significativas, bem como que não houve diferenças significativas entre as avaliações realizadas em dias sucessivos.

**Tabela 2.**

*Média e desvio-padrão das avaliações em escala Likert de rostos negros e brancos obtidas ao longo das medidas repetidas.*

Estímulos	Momento1	Momento 2	Momento 3
	Média (SD)	Média (SD)	Média (SD)
Faces Negras	2.84 (.72)	2.90 (.70)	2.8 (.80)
Faces Brancas	3.00 (.62)	3.10 (.72)	3.10 (.76)

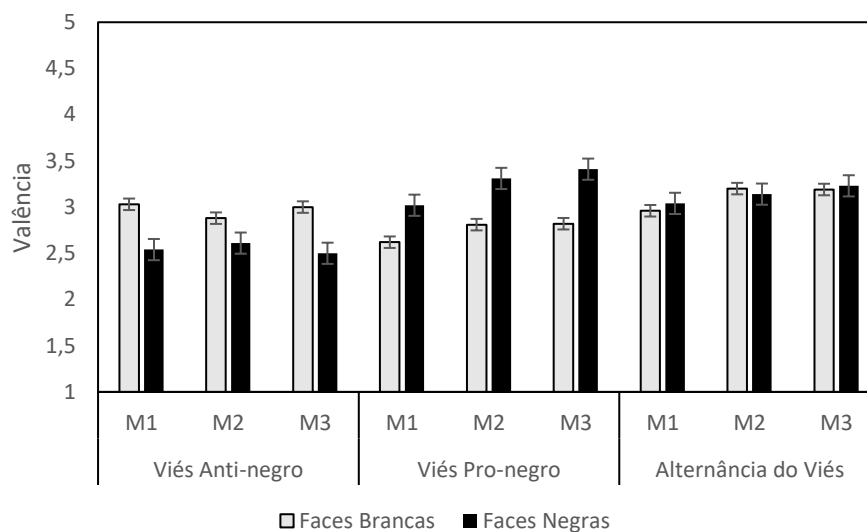
Uma análise individual indicou que 13 participantes avaliaram os rostos negros de forma mais positiva do que os rostos brancos nas três aplicações da medida, indicando um padrão de viés pró-negro. Dentro desse subgrupo, oito participantes eram pardos, três eram negros e dois eram brancos. Em contraste, sete participantes avaliaram os rostos brancos de forma mais positiva do que os rostos negros nos três dias de aplicação, sugerindo um padrão de viés antinegro. Nesse subgrupo, quatro participantes eram pardos, dois eram negros e um era branco. Por fim, 14 participantes apresentaram pequenas variações ao longo das repetições (alternando entre leve viés pró- e antinegro) nos diferentes dias de avaliação. Nesse subgrupo, nove participantes eram brancos, três eram pardos e um era negro.

A ANOVA de medidas repetidas do Tipo III, conduzida separadamente para cada subgrupo, não identificou diferenças significativas nas avaliações atribuídas a rostos negros e brancos. Adicionalmente, não foram observadas variações significativas ao longo dos três momentos de aplicação da medida, tampouco efeitos de interação significativos entre a avaliação de rostos negros e brancos e os dias de aplicação. Esses dados sugerem que, mesmo nos subgrupos que

apresentaram padrões leves de viés pró- ou antinegro, tais diferenças não se afastaram significativamente de zero. A Figura 1 ilustra as avaliações médias de valência atribuídas a rostos negros e brancos dentro desses subgrupos.

**Figura 1.**

*Valência média atribuída a rostos negros e brancos nos diferentes subgrupos.*



### Function Acquisition Speed Test (FAST)

De modo geral, os resultados indicaram uma taxa de aprendizagem mais elevada no bloco consistente em comparação ao bloco inconsistente ao longo dos três momentos de aplicação da medida (ver Tabela 2). Foi realizada uma ANOVA de medidas repetidas do Tipo III (Inclinação: bloco  $\times$  tempo) para avaliar diferenças nos resultados do FAST entre os blocos consistente e inconsistente nos três momentos de aplicação da medida (segunda-feira, quarta-feira e sexta-feira). A diferença na inclinação média de aprendizagem entre os blocos consistente e inconsistente ( $F(1) = 5,15$ ,  $p = 0,049$ ,  $\eta^2 = 0,002$ ) aproximou-se da significância, sugerindo um possível viés implícito antinegro na amostra. No entanto, testes *post hoc* pareados com correção de Bonferroni não indicaram diferenças significativas

no contraste entre os blocos consistente e inconsistente no primeiro ( $p = 0,22$ ), segundo ( $p = 0,12$ ) e terceiro ( $p = 0,18$ ) momentos de repetição da medida.

Não foram encontradas diferenças significativas entre os blocos em cada momento de aplicação da medida ( $F(2) = 0,22$ ,  $p = 0,80$ ,  $\eta^2 = 0,005$ ), nem efeito de interação entre bloco e tempo ( $F(2) = 0,07$ ,  $p = 0,93$ ,  $\eta^2 = 0,008$ ), indicando que a repetição não gerou mudanças significativas.

O desempenho do grupo indicou efeito FAST positivo nas três medidas basais, embora todos os valores tenham sido próximos de zero. O teste de Kruskal–Wallis não indicou diferenças significativas nos efeitos FAST ao longo das três medidas basais ( $H(2) = 0,28$ ,  $p = 0,86$ ), confirmando a ausência de mudanças no efeito FAST ao longo das repetições. A Tabela 3 apresenta as médias e os desvios-padrão das inclinações obtidas nos blocos consistente e inconsistente, bem como o efeito FAST médio em cada momento de avaliação.

### **Tabela 3.**

*Análise descritiva do desempenho no FAST ao longo dos diferentes dias de aplicação.*

Blocos	Momento 1 Média (SD)	Momento 2 Média (SD)	Momento 3 Média (SD)
Consistente	.23 (.04)	.25 (.05)	.25 (.06)
Inconsistente	.21 (.03)	.23 (.03)	.23 (.04)

Além disso, a análise individual indicou que 11 participantes (32,3%) apresentaram valores positivos consistentes do efeito FAST ao longo dos três momentos de avaliação, indicando aprendizagem mais rápida no bloco consistente e um possível viés implícito antinegro. Dentro desse subgrupo, seis participantes

eram pardos, três eram brancos e dois eram pretos. Para esse subgrupo, uma ANOVA de medidas repetidas revelou um efeito de interação significativo na inclinação dos blocos consistente e inconsistente,  $F(1) = 7,61$ ,  $p = 0,008$ ,  $\eta^2 = 0,03$ , indicando que a taxa de aprendizagem diferiu entre os blocos, embora com tamanho de efeito pequeno. No entanto, não houve efeito significativo dos três momentos de aplicação da medida ( $F(1) = 0,96$ ,  $p = 0,38$ ,  $\eta^2 = 0,001$ ), sugerindo que o efeito observado não variou significativamente entre os diferentes dias. Adicionalmente, não foi observada interação significativa entre bloco e momento de aplicação ( $F(2) = 0,33$ ,  $p = 0,71$ ,  $\eta^2 = 0,006$ ), indicando que a repetição da medida não influenciou significativamente o efeito da inclinação de aprendizagem entre os blocos.

A análise *post hoc* com correção de Bonferroni indicou um efeito principal da variável Bloco ( $p < 0,0001$ ), mas não do tempo de aplicação da medida ( $p > 0,47$ ). Contrastes pareados com correção de Bonferroni revelaram diferenças significativas entre os blocos consistente e inconsistente no primeiro ( $p = 0,005$ ), segundo ( $p = 0,0001$ ) e terceiro ( $p = 0,004$ ) dias de reaplicação.

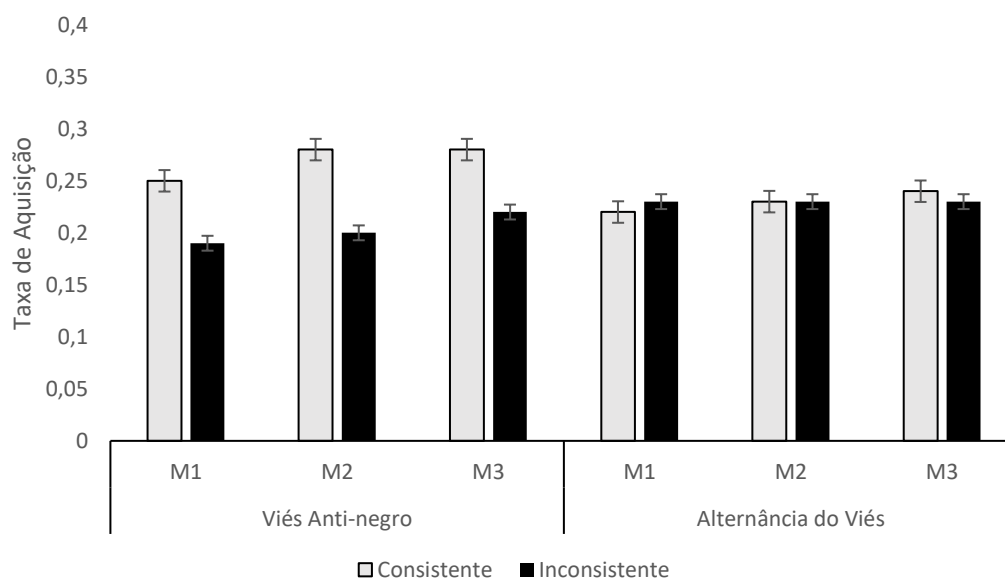
Adicionalmente, 23 participantes (67,6%) apresentaram pequenas variações no efeito FAST, alternando entre padrões de viés antinegro e pró-negro ao longo das repetições. Nesse subgrupo, doze participantes foram classificados como brancos, oito como pardos e cinco como negros.

A ANOVA de medidas repetidas do Tipo III não revelou diferenças estatisticamente significativas entre os blocos consistente e inconsistente  $F(1) = 0,23$ ,  $p = 0,95$ ,  $\eta^2 = 0,001$  entre os momentos de repetição da medida  $F(2) = 0,59$ ,  $p = 0,55$ ,  $\eta^2 = 0,003$ , nem interação entre bloco e momento de aplicação  $F(2) = 0,44$ ,  $p = 0,96$ ,  $\eta^2 = 0,002$ .

A Figura 2 apresenta o desempenho dos participantes nos blocos consistente e inconsistente, organizados em dois subgrupos: um que apresentou viés antinegro consistente ao longo das três aplicações do FAST e outro que alternou ligeiramente entre padrões de viés antinegro e pró-negro ao longo das aplicações repetidas da medida.

### Figura 2.

*Desempenho nos blocos consistente e inconsistente entre os subgrupos, destacando viés antinegro e variações no efeito FAST.*



### Correspondência entre as Medidas

Para a escala Likert, a valência atribuída aos rostos brancos foi subtraída da valência atribuída aos rostos negros, gerando um escore denominado “diferença”, que indica a tendência de viés pró- ou antinegro na amostra. Esse escore de diferença foi então correlacionado com o Efeito FAST, com o objetivo de explorar a convergência entre as medidas relacionadas ao viés apresentado pela

amostra. É importante destacar que, tanto para o Efeito FAST quanto para as diferenças na escala Likert, valores positivos indicam viés antinegro, enquanto valores negativos indicam viés pró-negro.

Os resultados do teste de Spearman revelaram correlações negativas (variando de  $-0,14$  a  $-0,20$ ), porém fracas e não significativas (todos  $p > 0,05$ ) nos três momentos de aplicação do instrumento. A Tabela 4 apresenta os coeficientes de correlação de Spearman ( $\rho$ ) e os respectivos valores de  $p$ , indicando a correlação entre o Efeito FAST e a escala Likert em cada momento de avaliação.

**Tabela 4.**

*Correlação entre o Efeito FAST e a Escala Likert em cada momento de aplicação da medida*

	Spearman $\rho$ (Rho)	p-valor
Momento 1	-.15	.41
Momento 2	-.20	.25
Momento 3	-.14	.44

Os 11 participantes que apresentaram viés antinegro no FAST foram posteriormente divididos em dois subgrupos distintos. Especificamente, sete participantes indicaram consistentemente um viés pró-negro na escala Likert ao longo dos três momentos de avaliação, ao mesmo tempo em que exibiram um Efeito FAST oposto, indicando viés antinegro nas três aplicações da medida. Assim, a relação entre rostos negros e estímulos positivos indicada pela medida explícita foi contradita pela relação entre rostos negros e estímulos negativos revelada pela medida implícita.

Os outros quatro participantes apresentaram consistentemente viés antinegro na escala Likert, evidenciando consistência entre as relações estímulo-resposta observadas nas medidas explícita e implícita. Uma limitação a considerar é que o estudo não controlou históricos individuais de exposição direta a situações de violência racial, variável que poderia exercer controle adicional sobre o responder avaliativo e contribuir para a variabilidade interindividual observada nos dados.

### **Discussão**

O presente estudo investigou a sensibilidade do FAST para identificar viés implícito em crianças, sua convergência com medidas de autorrelato do tipo Likert e a confiabilidade teste–reteste entre diferentes dias. Não foram observados indícios de viés racial explícito na escala de valência emocional. Contudo, o FAST indicou maior fluência no bloco consistente em relação ao inconsistente nos três momentos de avaliação, sugerindo viés implícito antinegro na amostra geral. Embora a ANOVA inicial tenha apontado possível diferença entre as inclinações dos blocos, análises post hoc não confirmaram esse efeito para a amostra total, mas revelaram diferença significativa em um subconjunto de 11 participantes. Não houve variações significativas no desempenho entre os momentos de aplicação, indicando estabilidade teste–reteste. Por fim, a correlação de Spearman foi negativa e não significativa, sugerindo ausência de convergência entre as medidas.

Pesquisas prévias da psicologia social experimental demonstram que mesmo crianças muito pequenas podem apresentar viés racial implícito (Baron & Banaji, 2006; Castelli et al., 2009; Cvencek et al., 2016; Dunham et al., 2007;

Hugenberg & Bodenhausen, 2004). No presente estudo, analisamos os efeitos do FAST ao longo de três aplicações com o objetivo de identificar participantes que apresentaram consistentemente viés antinegro ou pró-negro. Os resultados revelaram que 11 participantes exibiram consistentemente viés antinegro, demonstrando maior fluência no bloco consistente em todas as aplicações do FAST. A análise de variância indicou uma diferença significativa entre os blocos consistente e inconsistente para esses participantes, sugerindo taxas de aprendizagem mais rápidas na relação entre rostos negros e imagens negativas e entre rostos brancos e imagens positivas. Análises *post hoc* subsequentes corroboraram esses achados ao confirmar que a diferença significativa entre os blocos se manteve em cada momento de avaliação. Esses resultados estão em consonância com estudos anteriores que utilizaram o FAST para investigar vieses de gênero e racial, os quais também observaram taxas de aquisição mais rápidas de classes funcionais de resposta no bloco consistente com relações enviesadas (ver Cartwright et al., 2016; Rodrigues et al., 2020).

Além disso, entre as 11 crianças que apresentaram viés implícito antinegro, 72,72% foram classificadas como pardas ou negras. Esses achados dialogam com discussões da literatura social experimental, que indicam que percepções sociais sobre raça podem influenciar vieses dentro dos próprios grupos, resultando na emergência de viés racial antinegro em grupos oprimidos (ver Alexander & Carter, 2022; Araújo et al., 2022; Clark, 2004; Mizael et al., 2021; Tappan, 2006). De forma importante, a presença de viés racial na amostra parda também levanta a questão do colorismo, em que indivíduos com tons de pele mais claros podem apresentar viés antinegro direcionado a pessoas com tons de pele mais escuros

(Alexander & Carter, 2022; Clark, 2004; Landor, 2012). Uma limitação adicional é que o estudo não controlou em que medida essas crianças se reconheciam como pertencentes ao mesmo grupo racial das figuras apresentadas. A categorização racial não é um repertório natural ou espontâneo, mas produto de processos históricos de colonização que impõem ao sujeito uma identidade por diferenciação em relação ao colonizador (Fanon, 2020). Reconhecer-se como negro pode envolver repertórios de esquiva, identificação parcial ou ambivalência, dado que assumir rótulos historicamente associados a atributos negativos é um processo complexo e não-linear; a identidade negra se constitui por relações de oposição, sendo dependente da diferenciação em relação aos brancos e não sob controle dos eventos que deveriam, em condições outras, sustentar tal identificação (Gingles et al., 2022). O viés antinegro observado em crianças pardas e negras pode, portanto, refletir não apenas relações aprendidas sobre grupos externos, mas padrões de responder sob controle de contingências que dificultam a própria identificação com o grupo. Pesquisas futuras deveriam incluir avaliações do repertório de identificação racial dos participantes como variável de controle, bem como amostras mais amplas e diversas para melhor compreender essas dinâmicas na infância.

A tendência geral observada foi a de avaliar rostos negros de forma ligeiramente mais positiva do que rostos brancos na escala de valência emocional, uma medida explícita. Mesmo entre os participantes que avaliaram consistentemente rostos brancos de forma mais positiva ao longo das três repetições da escala, não houve diferença significativa entre a valência atribuída a rostos negros e brancos. Uma possível explicação para a ausência de viés

explícito na amostra é a influência de contingências escolares e familiares que promovem explicitamente valores de igualdade racial, como atividades pedagógicas voltadas à representatividade e à consciência racial, que podem ter estabelecido regras verbais contrárias à expressão explícita de avaliações negativas de pessoas negras (ver Sawyer & Gampa, 2018). No entanto, a análise de correlação indicou ausência de convergência entre as medidas implícitas e explícitas. Esse é um achado comum na literatura social, uma vez que vieses implícitos e explícitos não são necessariamente esperados correlacionar-se (e.g., Cartwright et al., 2016; Dasgupta & Greenwald, 2001; Fazio & Olson, 2004; Hofmann et al., 2005; Greenwald et al., 2009; Payne, 2001; Rodrigues et al., 2022; Sawyer & Gampa, 2018).

A medida explícita envolve respostas relacionais complexas, sem restrição temporal, nas quais os indivíduos podem ser influenciados por diversas contingências, como punição por expressar visões preconceituosas ou reforçamento por emitir respostas socialmente desejáveis. Embora se esperasse que crianças apresentassem maior correspondência entre medidas implícitas e explícitas do que adultos, em função de uma maior probabilidade de expressar vieses na escala Likert (Levy & Hughes, 2009), os resultados do presente estudo são consistentes com achados prévios em amostras adultas (ver Cartwright et al., 2016; Rodrigues et al., 2020). Ademais, embora não estatisticamente significativa, a correlação negativa entre a escala de valência emocional e o FAST sugere uma tendência à emissão de respostas contrastantes. De fato, a maioria das crianças que apresentou viés explícito pró-negro na escala Likert exibiu resultados opostos no FAST. Esses achados indicam a possibilidade de que algumas crianças nessa

faixa etária já estejam ocultando vieses identificados em respostas relacionais breves, de modo a emitir respostas socialmente aceitas em medidas explícitas (ver Van de Mortel, 2008; Krumpal, 2013). Esse aspecto constitui um campo promissor para investigações futuras.

Neste estudo, também investigamos a confiabilidade teste–reteste do FAST e da escala Likert ao longo de aplicações repetidas (Jorm et al., 1989; Kalra et al., 2019; Schultz et al., 2004). Os resultados não indicaram efeitos de interação significativos entre os blocos ao longo das repetições da medida nos três momentos avaliados, nem mesmo no subgrupo que apresentou viés racial antinegro. Além disso, não foram observadas diferenças significativas no efeito FAST ao longo dos três dias de aplicação. De modo semelhante, a escala Likert não apresentou mudanças significativas entre as aplicações. Assim, ambas as medidas parecem passíveis de uso como medidas repetidas, ampliando sua utilidade para a identificação do preconceito racial e para a avaliação da eficácia de intervenções.

No entanto, é importante destacar que este estudo não exclui a possibilidade de efeitos de prática decorrentes de exposições repetidas. Enquanto a confiabilidade teste–reteste avalia a estabilidade dos resultados ao longo do tempo, os efeitos de prática envolvem melhorias de desempenho resultantes da repetição, podendo afetar a validade do instrumento como medida (Jones, 2015). Por exemplo, um estudo conduzido por Lai et al. (2014) demonstrou que a exposição repetida a blocos inconsistentes consecutivos no IAT resultou em melhoria de desempenho, sendo essa abordagem utilizada como intervenção para reduzir o viés racial implícito (ver também Lai et al., 2016). Nossos achados

sugerem que a fluência no FAST, quando aplicada em dias distintos, não apresenta mudanças significativas, mesmo em subgrupos que alternaram entre leve viés pró- e antinegro ao longo das repetições. Esses resultados têm implicações para estudos que utilizam intervalos espaçados entre avaliações de viés, aplicações de intervenção e pós-testes (e.g., Mizael et al., 2016, 2020), permitindo comparações com aplicações pós-intervenção.

É fundamental reconhecer a importância de pesquisas voltadas à identificação do viés racial implícito, especialmente em crianças, pois elas oferecem subsídios valiosos para compreender a organização do ambiente verbal em contextos como escolas, famílias e comunidades (Aboud, 1993; Baron & Banaji, 2006; Clark & Clark, 1947; Dasgupta & Greenwald, 2001; Levy & Hughes, 2009). A compreensão dessas dinâmicas também contribui para o enfrentamento não apenas de preconceitos individuais, mas de vieses sistêmicos, como o racismo institucional, possibilitando o desenvolvimento de estratégias educacionais culturalmente sensíveis e intervenções voltadas à promoção de ambientes mais inclusivos desde a infância. Ademais, a ausência de correspondência entre medidas implícitas e explícitas observada neste estudo reforça a complexidade do preconceito racial, frequentemente não relatado de forma direta (Sawyer & Gampa, 2018), indicando a necessidade de estratégias que abordem formas indiretas de viés. Esses achados podem orientar a seleção de intervenções individualizadas, direcionadas tanto ao preconceito explícito (e.g., Mizael et al., 2016, 2020) quanto ao implícito (e.g., Gonzalez et al., 2017; Lai et al., 2014, 2016; Sinclair et al., 2005).

O uso do FAST como medida em estudos voltados à identificação de viés implícito em crianças representa uma via promissora para a pesquisa social no âmbito da análise do comportamento. Trata-se de um método simples e rápido para avaliar o viés racial. Ao introduzir o FAST como medida repetida, o presente estudo demonstrou a confiabilidade teste–reteste de uma medida comportamental implícita ao longo de exposições repetidas em dias distintos. O FAST empregado neste estudo mostra-se potencialmente viável como uma ferramenta confiável para identificar viés racial e para uso em medidas repetidas, mantendo sensibilidade à organização de relações verbais socialmente estabelecidas. Contudo, é essencial reconhecer limitações inerentes, incluindo o tamanho amostral reduzido, decorrente dos desafios de conduzir pesquisas com crianças. Outro aspecto relevante é a utilização ainda restrita do FAST no contexto da pesquisa social, uma vez que apenas um número limitado de estudos o empregou para identificar vieses (ver Cartwright et al., 2016; Rodrigues et al., 2022). Assim, o FAST constitui atualmente uma metodologia ainda pouco explorada para a avaliação de relações verbais em domínios sociais sensíveis, sendo necessários novos estudos para explorar plenamente seu potencial de contribuição para o avanço da pesquisa social no campo da análise do comportamento.

-----

## **Comentário**

### **Estabilidade sem convergência: Estamos diante de um problema?**

O conjunto de dados apresentado neste estudo não responde a todas as perguntas que motivaram sua realização, e esse é precisamente o seu valor científico. Ao examinar medidas explícitas e implícitas de viés racial antinegro em crianças ao longo de aplicações repetidas, este trabalho contribui menos por demonstrar efeitos robustos e mais por esclarecer quais perguntas podem ser feitas de forma responsável a partir dessas medidas.

Avaliações explícitas produziram padrões pouco diferenciados entre estímulos raciais e mostraram alta estabilidade ao longo do tempo. Quando considerados em conjunto com os resultados do Teste da Boneca, esses dados apontam para forte controle por contingências de deseabilidade social. Um exemplo reside nas falas emitidas pelas crianças durante o teste, como “não podemos ser racistas” ou “eu não tenho preconceito”. Esse controle tende a reduzir a sensibilidade dessas medidas às relações avaliativas efetivamente estabelecidas, especialmente quando comparadas a medidas sob maior pressão temporal ou menor controle verbal.

No entanto, a presença desse controle não deve ser interpretada exclusivamente como uma limitação metodológica. O fato de as crianças demonstrarem conhecimento explícito de que o preconceito é socialmente reprovável e evitarem responder de forma preconceituosa quando dispõem de tempo para elaborar suas respostas constitui, em si, um

dato relevante. Esse padrão sugere a aprendizagem de normas sociais antirracistas e pode refletir práticas educativas e institucionais voltadas à condenação do preconceito racial.

Nesse sentido, a desejabilidade social indica não apenas restrições às medidas explícitas, mas também avanços no controle social exercido sobre o responder racialmente enviesado. A dissociação observada entre respostas explícitas e implícitas não aponta, portanto, para ausência de aprendizagem antirracista, mas para a coexistência de diferentes níveis de controle, normativo, relacional e contextual, sobre o comportamento. Essa constatação desloca a questão empírica central da simples detecção do viés para a análise de quando, como e sob quais condições diferentes formas de controle prevalecem ao longo do desenvolvimento.

Em contraste, o FAST mostrou-se capaz de identificar diferenças consistentes na forma como parte das crianças relacionava estímulos raciais a conteúdos positivos e negativos, mesmo quando essas relações não eram expressas de forma direta. Além disso, essas diferenças permaneceram estáveis ao longo das reaplicações da tarefa, sem indícios de que a repetição, por si só, alterasse o desempenho. Essa combinação, sensibilidade para captar padrões avaliativos e estabilidade ao longo do tempo, indica que o FAST pode ser utilizado como uma medida confiável em estudos longitudinais e em pesquisas que investiguem os efeitos de intervenções sobre o viés racial.

Este estudo também esclarece um aspecto metodológico frequentemente negligenciado: a necessidade de examinar a estabilidade das próprias medidas antes de utilizá-las como critérios de mudança. Ao demonstrar que a repetição, quando espaçada, não produziu alterações sistemáticas nas respostas, cria-se uma base empírica para a interpretação de efeitos interventivos futuros. Sem esse passo, qualquer mudança observada

poderia ser atribuída ao simples contato repetido com o procedimento de medida, e não à reorganização funcional das relações avaliativas.

Os resultados mostraram ausência de convergência entre os resultados do FAST e das avaliações explícitas na escala de valência emocional. No entanto, a ausência de correspondência sistemática entre as medidas explícitas e implícitas não deve ser tratada como uma limitação instrumental, mas como uma propriedade do fenômeno sob investigação. O viés racial não se manifesta como um repertório unitário, acessível por uma única classe de respostas. Ele emerge de formas distintas sob diferentes arranjos de contingências, exigindo da pesquisadora cautela ao inferir sua presença ou ausência com base em um único indicador.

Esse achado tem implicações diretas para o planejamento experimental. Quando uma medida explícita se mantém estável, mas tende a refletir respostas socialmente desejáveis, e uma medida implícita revela padrões avaliativos distintos, por vezes opostos ou sensíveis ao contexto, fica claro que elas não estão medindo a mesma coisa. Nessa condição, a questão central deixa de ser qual medida estaria “correta” e passa a ser que tipo de informação cada uma torna visível e, de forma igualmente importante, o que cada uma tende a mascarar. Medir viés racial, portanto, exige reconhecer que nenhuma medida, isoladamente, é suficiente para descrever o fenômeno, ainda que algumas sejam mais informativas do que outras, dependendo das condições sob as quais são aplicadas.

É nesse ponto que se impõe uma ampliação do escopo metodológico. Se diferentes medidas capturam diferentes aspectos do viés racial, algumas mais sensíveis ao controle temporal, outras mais permeáveis a contingências sociais explícitas, então a investigação empírica precisa avançar para arranjos multimodais, capazes de integrar essas fontes de variação em um mesmo delineamento. Não se trata de acumular instrumentos, mas de

orquestrar medidas que permitam observar o fenômeno sob diferentes óticas, em vez de forçá-lo a caber em uma única forma de mensuração.

O próximo estudo empírico que compõe este capítulo parte exatamente desse pressuposto. Em vez de perguntar se uma medida é superior a outra, ele investiga como diferentes tipos de medidas, relacionais, avaliativas e naturalísticas, convergem, divergem e se organizam funcionalmente quando aplicadas ao mesmo conjunto de participantes. Ao fazê-lo, desloca o foco da detecção isolada do viés racial para a compreensão de sua estrutura funcional, criando as condições necessárias para intervenções que não dependam de um único indicador, mas de padrões consistentes de responder ao longo de múltiplos contextos.

Como Sidman (1994) já advertia, avanços conceituais duradouros não decorrem de respostas apressadas, mas da disposição em permitir que os dados nos digam que tipo de pergunta ainda não sabemos formular corretamente. Este estudo cumpre esse papel. O próximo assume a tarefa mais difícil: responder às perguntas que este tornou inevitáveis.

\*\*\*\*\*

### **Estudo 3 - Identificação do preconceito racial em crianças com uso de diferentes medidas: Uma análise exploratória multimodal<sup>7</sup>**

A identificação do viés racial na infância requer medidas capazes de discriminar diferentes classes de respostas sob controle de estímulos raciais e variáveis contextuais. Embora a literatura analítico-comportamental tenha avançado na mensuração de repertórios avaliativos raciais, poucos estudos investigaram sistematicamente a sensibilidade de diferentes medidas e as interações contextuais. O presente estudo examinou a sensibilidade medidas explícitas e implícita na identificação de respostas de preconceito racial em crianças, considerando a interação entre raça e gênero do estímulo e a raça do participante. Participaram 80 crianças (M = 8,5 anos), avaliadas em um delineamento transversal com medidas afetivas explícitas (*Self-Assessment Manikin – SAM*), escolhas sociais e desempenho relacional implícito (*Function Acquisition Speed Test – FAST*). Os resultados indicaram efeitos principais robustos do gênero da face sobre valência, dominância e ativação emocional, além de interações significativas entre raça e gênero, particularmente para faces femininas. As escolhas sociais variaram conforme o contexto da tarefa, com maior seleção de estímulos brancos em atividades acadêmicas e de estímulos negros em contextos associados à atividade física. No FAST, observou-se uma interação entre raça do participante e tipo de bloco, com crianças brancas apresentando maior fluência em relações consistentes com o viés racial dominante e crianças pardas em relações inconsistentes. Em conjunto, os achados indicam que o viés racial infantil envolve repertórios múltiplos, funcionalmente distintos e modulados por contingências contextuais, reforçando a necessidade de abordagens multimodais.

---

<sup>7</sup> Passarelli, D. A., Mizael, T. M., Roche, B., & De Rose, J. C. C. (2026). *Identificação do preconceito racial em crianças com uso de diferentes medidas: Uma análise exploratória multimodal* [manuscrito em preparação a ser submetido a periódico científico].

## Introdução

Do ponto de vista da Análise do Comportamento, o preconceito racial tem sido investigado como atitudes negativas aprendidas em relação a membros de determinados grupos étnico-raciais, entendidas funcionalmente como padrões de responder avaliativo sob controle de estímulos raciais (Mizael & de Rose, 2017; Mizael et al., 2016; Passarelli, Mizael et al., 2024). Especificamente, o preconceito racial antinegro refere-se a avaliações negativas dirigidas a indivíduos socialmente categorizados como negros, independentemente de sua origem nacional ou étnica (Corrington et al., 2023; Passarelli, Mizael, 2024).

No campo experimental, essas atitudes têm sido operacionalizadas como *avaliações*, definição amplamente compartilhada com a psicologia social experimental e que permite sua mensuração sistemática em contextos controlados (Gawronski, 2007; Hayes et al., 2001). Cabe destacar que, na Análise do Comportamento, tais avaliações são compreendidas como responder relacional, cuja topografia e função podem variar em função das contingências e dos controles contextuais envolvidos (Barnes-Holmes et al., 2010; Guerin, 2005; Kohlenberg, 1991; Passarelli, Roche et al., 2025). Avaliações 'explícitas' tendem a envolver responder relacional estendido e verbalmente mediado, emitido sem restrição temporal. Avaliações implícitas, por sua vez, são capturadas por meio de propriedades do responder, como fluência, latência ou velocidade de aquisição sob contingências específicas, que escapam ao controle verbal direto do participante (ver Barnes-Holmes et al., 2010; Passarelli, Mizael et al., 2024; Watters et al., 2023). Assim, atitudes explícitas e implícitas não são tratadas como entidades distintas, mas como topografias diferentes de um mesmo processo

funcional, diferenciadas principalmente pelo grau de mediação verbal e pelo controle temporal da resposta.

Nesse contexto, o Functional Acquisition Speed Test (FAST) tem sido proposto como uma medida implícita sensível à história de aprendizagem relacional dos indivíduos (Cartwright et al., 2016; Passarelli, Mizael et al., 2024; Rodrigues et al., 2022; Roche et al., 2025). Evidências empíricas recentes demonstraram que o FAST é capaz de detectar avaliações negativas implícitas relacionadas a estímulos raciais, bem como de apresentar estabilidade ao longo de aplicações repetidas, sugerindo sensibilidade e consistência do instrumento para investigar repertórios avaliativos raciais (Passarelli, Mizael et al., 2024).

Além disso, o FAST tem se mostrado uma ferramenta socialmente sensível para detectar padrões de responder relacional, utilizando variáveis sociodemográficas como gênero e raça como fatores. A medida obtida não é uma avaliação direta dos estímulos, mas a velocidade de aquisição de relações condicionais sob contingências específicas, propriedade do responder que reflete a força de redes relacionais pré-existentes (O'Reilly et al., 2012; Roche et al., 2025). Em Roche et al. (2025), por exemplo, o gênero do participante emergiu como fator significativo, de modo que mulheres apresentaram maior velocidade de aquisição de relações condicionais envolvendo faces femininas e estímulos positivos quando comparadas a homens, padrão interpretado como consistente com viés favorável a membros do próprio grupo (ver Experimento I). De forma semelhante, um segundo experimento envolvendo avaliação de faces negras revelou viés antinegro significativo apenas em participantes brancos; participantes não brancos, embora com tendência na mesma direção, não apresentaram

diferença estatisticamente significativa entre os blocos, resultado atribuído pelos autores à composição heterogênea desse subgrupo, predominantemente de background misto, o que pode ter introduzido variabilidade relacionada ao repertório de identificação racial dos participantes (Roche et al., 2025).

Apesar desses avanços, a literatura empírica sobre medidas de preconceito racial tem se concentrado predominantemente na dimensão avaliativa, entendida como atribuição explícita ou implícita de valência positiva ou negativa a estímulos raciais (Mizael & de Rose, 2016; Mizael et al., 2016, 2020; Passarelli, Mizael, et al., 2024; Roche et al., 2024). No entanto, estudos recentes têm destacado que o que é genericamente denominado “preconceito racial” envolve múltiplas classes de respostas que não são necessariamente funcionalmente equivalentes (ver Passarelli, Roche et al., 2024; Passarelli, Mizael & de Rose, 2023). Avaliações afetivas, preferências sociais e respostas de discriminação racial podem estar sob controle de variáveis antecedentes e consequentes distintas, não apresentando convergência sistemática entre si (e.g., Dasgupta & Greenwald, 2001; Greenwald et al., 1998; Payne, 2001; Payne et al., 2008; Sawyer & Gampa, 2018). Essa dissociação não decorre de uma oposição entre medidas implícitas e explícitas, tampouco de supostos processos cognitivos distintos, mas de variações no controle de estímulos e nas contingências contextuais sob as quais essas respostas são emitidas (ver Passarelli, Roche, Mizael et al., submetido). Assim, diferentes repertórios racialmente relevantes podem ser selecionados e mantidos em função das demandas específicas de cada contexto.

Por exemplo, membros de um mesmo grupo étnico-racial podem ser avaliados negativamente em determinados contextos funcionais (e.g., atividades

acadêmicas ou intelectuais) e positivamente em outros (e.g., esportes ou atividades físicas). Ainda que topograficamente podem parecer “positivas”, no contexto do preconceito racial, essas avaliações podem operar de forma restritiva, limitando oportunidades, expectativas e formas de participação social atribuídas a indivíduos com base na cor da pele (Conrado & Ribeiro, 2017). Nesse sentido, o caráter prejudicial do preconceito racial não reside apenas na presença de avaliações negativas globais, mas na organização diferencial de contingências que restringem o acesso a reforçadores sociais e institucionais (Mizael & Sampaio, 2019; Saini & Vance, 2020; Passarelli et al., 2023). Dessa forma, a investigação empírica do preconceito racial demanda abordagens multimétodo capazes de capturar diferentes facetas do responder racialmente relevante, incluindo avaliações afetivas, padrões de escolha social e desempenho em tarefas de aprendizagem relacional.

Entre essas facetas, destacam-se não apenas avaliações de valência, mas também relações de hierarquia e poder (e.g., dominância), níveis de ativação emocional (e.g., excitação fisiológica associada a estados de calma versus agitação ou medo), empatia ou tomada de perspectiva (ver Suarez et al., 2024) e padrões de preferência e discriminação (ver Correll et al., 2002) em contextos sociais concretos.

Diante disso, o presente estudo parte desses pressupostos para expandir achados anteriores (Passarelli, Mizael et al., 2024; Roche et al., 2025) e investigar de forma mais abrangente o responder relacional vinculado à raça em crianças. Especificamente, este estudo teve como objetivos: (a) examinar se as avaliações afetivas explícitas na Escala Self-Assessment Manikin (SAM; Bradley & Lang,

1994) variam em função da raça e do gênero dos estímulos faciais, considerando as dimensões de valência, dominância e ativação emocional; (b) investigar se a raça dos participantes atua como variável contextual que modula essas avaliações, tendo em vista que o repertório de identificação racial pode variar entre indivíduos pertencentes a um mesmo grupo fenotípico; (c) testar a generalidade do FAST com o uso de palavras avaliativas estereotípicas em uma amostra infantil, avaliando diferenças de desempenho entre blocos consistentes e inconsistentes em função da raça das crianças; (d) analisar padrões de preferência e discriminação social por meio de um Teste da Boneca modificado, composto por cinco bonecas com diferentes tonalidades de pele, permitindo examinar efeitos de colorismo em situações sociais e acadêmicas variadas; e (e) avaliar em que medida essas diferentes medidas capturam formas distintas de expressão do racismo, contribuindo para a compreensão da validade convergente e discriminante dos instrumentos utilizados.

Ao integrar diferentes medidas explícitas e implícitas, bem como variáveis intra-sujeitos (raça e gênero dos estímulos) e inter-sujeitos (raça dos participantes), o presente estudo busca ampliar a compreensão funcional do preconceito racial na infância. Essa abordagem permite examinar como diferentes classes de respostas, avaliativas, afetivas, relacionais e de escolha social, podem ser selecionadas ou moduladas por contextos específicos, revelando padrões de responder racialmente relevante que não seriam detectados por medidas isoladas

## Método

### Participantes e Local

Participaram do estudo 80 crianças ( $M = 8,5$  anos,  $DP = 0,5$ ), estudantes do 3º e 4º anos do Ensino Fundamental I de uma escola pública estadual localizada no interior do estado de São Paulo. A amostra foi composta por 51 participantes do gênero feminino (63,75%) e 29 do gênero masculino (36,25%).

A classificação racial seguiu as categorias do Instituto Brasileiro de Geografia e Estatística (IBGE) e foi declarada pelos responsáveis legais no Termo de Consentimento Livre e Esclarecido (TCLE). Com base nessa classificação, 43 crianças (53,5%) foram identificadas como brancas e 37 crianças (46,5%) como pardas, não havendo participantes classificados como pretos.

### Instrumentos

#### Self-Assessment Manikin

O *Self-Assessment Manikin* (SAM; Bradley & Lang, 1994): foi utilizado como medida explícita multidimensional para avaliar respostas afetivas às imagens de quatro faces negras (2 femininas e 2 masculinas) e quatro faces brancas (2 femininas e 2 Masculinas). Foram avaliadas três dimensões afetivas independentes: valência, excitação (*arousal*) e dominância, que são domínios já presentes na escala.

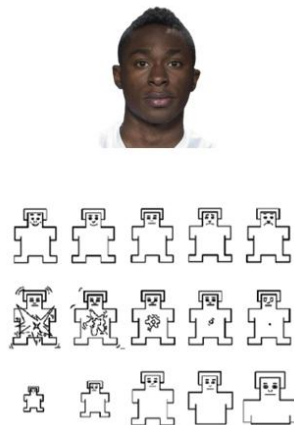
Cada dimensão foi avaliada em uma escala pictórica de 5 pontos, baseada em figuras humanas esquemáticas. Para dimensão valência, 1-2 pontos indicaram valência negativa; 3 valências neutra; 4-5 valência positiva. Para a dimensão dominância, 1-2 indicava dominado; 3 nem dominante nem dominado; 4-5 dominante. Importante que a dimensão *arousal*, menor pontuação indicou maior

ativação emocional, indicando respostas de agitação (1-2), 3 nem agitado nem calmo; enquanto maior pontuação indicou menor ativação emocional (4-5), indicando calma. Essa inversão foi feita para que ficasse compatível com as outras escalas, onde menores valores indicavam respostas desfavoráveis em relação às faces negras.

As imagens de rostos foram recuperadas do *FaceResearch* ([faceresearch.org](http://faceresearch.org)), contendo duas faces femininas negras, duas faces femininas brancas, duas faces masculinas negras e duas faces masculinas brancas. As faces foram apresentadas individualmente no centro da tela, com equilíbrio quanto ao gênero e raça e expressão facial neutra. Para cada dimensão, os participantes indicavam sua resposta selecionando a figura correspondente ao ponto da escala que melhor representava sua avaliação do estímulo apresentado.

### Figura 1.

*Escala SAM apresentada aos participantes com exemplo de face apresentada*



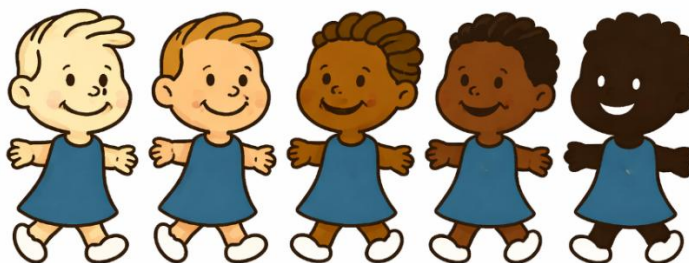
*Nota.* Foram apresentadas total de oito faces, sendo quatro negras e quatro brancas, cada face apresentada em tentativas separadas. As três dimensões foram colocadas abaixo da foto e uma avaliada de cada vez.

## Teste da Boneca

O Teste da Boneca Modificado consistiu em uma tarefa computadorizada de preferência social, inspirada no paradigma clássico de Clark e Clark (1947). O conjunto de estímulos foi apresentado em tela de computador e consistiu no desenho de cinco bonecas fisicamente idênticas, variando exclusivamente quanto ao tom de pele, organizadas em uma gradação contínua do mais claro ao mais escuro. No teste tradicional de Clark & Clark (1947), eram apresentadas à criança duas bonecas tridimensionais — uma negra e uma branca —, havendo perguntas acerca de atributos avaliativos (e.g., boa, má, bonita, feia), bem como questões de identificação racial. A Figura 2 apresenta os estímulos utilizados no presente estudo em substituição das duas bonecas tridimensionais do teste clássico.

### Figura 2.

*Estímulos utilizados no Teste da Boneca Modificado*



Em cada tentativa da tarefa computadorizada, uma pergunta de preferência social era exibida na parte superior da tela e os cinco estímulos eram apresentados simultaneamente na parte inferior. A criança respondia apontando diretamente para a tela do computador as bonecas que desejava selecionar. As perguntas avaliativas do teste original foram substituídas por perguntas de

preferência em contextos de atividades sociais corriqueiras (e.g., "Escolheria alguma para ser sua amiga?", "Escolheria alguma para fazer um trabalho para nota?", "Escolheria alguma para o time da Educação Física?", "Escolheria alguma para o seu grupo?"). Em cada pergunta, as crianças podiam selecionar livremente quantas bonecas desejassem. Essa modificação teve como objetivo identificar padrões de preferência racial em contextos socialmente relevantes, aproximando o procedimento de situações análogas às da vida cotidiana.

As respostas foram registradas pelo número da boneca selecionada (de 1 a 5, correspondendo à gradação do tom de pele mais claro ao mais escuro) e, posteriormente, convertidas em categoria binária (Negra / Branca) para fins de análise. O formato computadorizado permitiu o registro preciso das escolhas sem exigir verbalizações avaliativas diretas, reduzindo as demandas verbais durante a tarefa.

### ***Functional Acquisition Speed Test (FAST)***

O *Functional Acquisition Speed Test* – versão modificado (FAST; O'Reilly et al., 2012; Cartwright et al., 2016; Passarelli, Mizael, et al., 2024) foi utilizado como medida implícita de preconceito racial, avaliando a fluência de aquisição de relações funcionais entre estímulos raciais e estímulos avaliativos.

O protocolo do FAST foi similar ao empregado por Passarelli, Mizael et al. (2024), incluindo um bloco de prática (20 tentativas) com estímulos não relacionados ao objetivo do estudo, seguido de dois blocos de teste (50 tentativas cada), um consistente e outro inconsistente com o viés racial socialmente dominante. As instruções fornecidas às crianças foram mínimas, limitando-se à indicação das teclas de resposta e à orientação para responder o mais

rapidamente possível, com base no feedback apresentado, sem explicitação das relações avaliativas envolvidas.

Os estímulos-alvo foram as mesmas oito faces negras e brancas utilizadas na escala SAM (quatro negras e quatro brancas) e palavras avaliativas escritas, uma vez que as crianças apresentavam fluência leitora compatível com a tarefa. As palavras de valência positiva incluíram bom, inteligente, bonito e calmo, enquanto as palavras de valência negativa incluíram mau, feio, agressivo e burro. Foram apresentadas 50 tentativas em cada bloco (consistente e inconsistente). As respostas eram emitidas por meio das teclas Z e M do teclado, com feedback corretivo imediato ("correto" ou "incorreto") após cada tentativa.

O desempenho no FAST foi quantificado por meio da inclinação da curva cumulativa de respostas corretas em função do tempo, refletindo a taxa de aquisição das classes funcionais de resposta em cada bloco. Quanto maior a inclinação, maior a fluência de aprendizado, indicando aquisição mais rápida das relações condicionais treinadas. As inclinações dos blocos consistente e inconsistente foram analisadas separadamente, sem subtração entre blocos.

### **Procedimento**

O procedimento foi aprovado pelo Comitê de Ética em Pesquisa da Universidade Federal de São Carlos (CAAE nº 48028721.0.0000.5504) e pela Secretária Estadual de Educação do Estado de São Paulo. A participação foi voluntária, com consentimento dos responsáveis legais e assentimento das crianças, sendo os estudos conduzidos mediante autorização da Secretaria Pública de Educação e da direção das escolas envolvidas.

O estudo adotou um delineamento transversal, no qual cada participante foi avaliado em uma única sessão. As coletas foram realizadas individualmente, em ambiente escolar reservado, silencioso e com condições adequadas para a realização das tarefas, contando apenas com a presença da pesquisadora e da criança.

Cada sessão teve duração aproximada de 20 a 25 minutos e seguiu uma ordem fixa de aplicação dos instrumentos, adotada para minimizar possíveis efeitos de exposição das medidas implícitas sobre as respostas explícitas. A sequência foi: *Self-Assessment Manikin* (SAM), Teste da Boneca e, por fim, o *Function Acquisition Speed Test* (FAST).

Inicialmente, foi aplicado o *Self-Assessment Manikin* (SAM). As crianças avaliaram imagens de faces negras e brancas em três dimensões afetivas: valência, ativação (arousal) e dominância. Em cada tentativa, a face era apresentada por aproximadamente dez segundos, acompanhada das três escalas do SAM. Embora todas as escalas permanecessem visíveis, a avaliação era conduzida de forma sequencial, com a criança respondendo uma dimensão por vez, conforme a instrução da pesquisadora.

Para a dimensão de valência, a pesquisadora explicou que a criança poderia indicar o quanto a face parecia boa ou ruim. Avaliações positivas foram descritas como associadas a coisas boas, agradáveis ou que deixariam a criança feliz. A pesquisadora orientava o participante a observar que, nos escores 4 e 5, a figura apresentava expressão sorridente, sendo 5 correspondentes a “muito bom/positivo” (sorriso mais intenso) e 4 a “bom/positivo” (sorriso menos intenso). Em seguida, a pesquisadora explicava que a face também poderia ser avaliada

como negativa, isto é, ruim, desagradável ou capaz de deixá-lo triste. Para isso, eram apresentadas as figuras correspondentes aos escores 1 e 2, ambos com expressões tristes, sendo 1 equivalente a “muito ruim/negativo” e 2 a “ruim/negativo”. Por fim, foi explicado que a face poderia ser considerada neutra, isto é, nem boa nem ruim, situação representada pelo escore 3, correspondente a uma expressão intermediária do manequim.

Para a dimensão de *arousal*, a pesquisadora explicou que a criança deveria avaliar o quanto a face parecia deixar a pessoa calma ou com medo. Foi dito que algumas pessoas, ao verem certos rostos, podem se sentir tranquilas (como se estivessem vendo alguém familiar que deixa a gente calmo e seguro), enquanto outras podem se sentir assustadas, tensas ou em alerta (como se estivessem vendo alguém perigoso). A pesquisadora orientava a criança a observar os manequins e explicou que os números 4 e 5 representavam situações em que a face o deixava calmo. O escore 5 significava que a face parecia muito calma e não dava medo algum, enquanto o escore 4 indicava que a face parecia calma. Em contraste, os escores 1 e 2 representavam situações em que a face causava medo ou deixava a pessoa assustada. O escore 1 indicava que a face parecia dar muito medo, enquanto o escore 2 indicava que a face parecia dar um pouco de medo. O escore 3 era utilizado quando a criança achava que a face não era nem calma nem assustadora.

Para a dimensão de dominância, a pesquisadora explicou que a criança deveria avaliar o quanto a face parecia dominante ou submissa. Escores mais altos indicavam que a face parecia mais forte, confiante ou no controle, enquanto escores mais baixos indicavam que a face parecia mais fraca, submissa ou

insegura. A pesquisadora orientava o participante a observar que, nos escores 4 e 5, o manequim apresentava uma postura associada a maior dominância, sendo 5 correspondentes a “muito dominante” e 4 a “dominante”. Por outro lado, os escores 1 e 2 representavam menor dominância, com manequins indicando submissão ou falta de controle, sendo 1 equivalente a “muito submisso” e 2 a “submisso”. O escore 3 indicava uma avaliação intermediária, na qual a face não parecia nem dominante nem submissa.

Em seguida, foi administrado o Teste da Boneca, conduzido por meio da apresentação de um cenário imaginativo padronizado. As crianças foram instruídas a imaginar que as bonecas representavam novos alunos que haviam acabado de chegar à sua sala de aula. A pesquisadora solicitava que a criança observasse atentamente todas as bonecas antes de responder às perguntas. A partir desse cenário, foram feitas perguntas específicas relacionadas a preferências em diferentes contextos sociais e acadêmicos, tais como: 1) alguma você escolheria para ser sua amiga? 2) alguma você escolheria para atividade para nota? 3) alguma escolheria para seu time da educação física? 4) alguma é sua preferida? Em cada pergunta, a criança realizava uma escolha livre, sem restrição de tempo e sem qualquer tipo de feedback por parte da pesquisadora. Cada pergunta foi tratada como uma situação independente, de modo a evocar padrões de escolha sob diferentes controles contextuais, permitindo a análise de preferências e possíveis respostas de discriminação em função do contexto da tarefa.

Por fim, foi aplicado FAST, conforme protocolo descrito anteriormente em Passarelli, Mizael et al., (2024). As instruções fornecidas foram mínimas e

padronizadas de estudos anteriores, restringindo-se à indicação das teclas de resposta e à orientação para responder com base no feedback apresentado pelo programa. Cada tentativa foi seguida de feedback corretivo imediato (“correto” ou “incorreto”), conforme as contingências programadas para cada bloco.

Participantes iniciaram pelo bloco de prática, com estímulos não relacionados ao estudo, para gerar familiaridade com a tarefa. Já o bloco de testes foi randomizado, de modo que metade dos participantes começaram pelo bloco consistente e a outra metade pelo bloco inconsistente. Ao término da sessão, as crianças foram informadas de que a atividade havia sido concluída e agradecidas pela participação. Uma sessão adicional com devolutiva e psicoeducação sobre preconceito racial foi ministrada para cada participante.

### **Análise de Dados**

Uma análise de poder estatístico *a posteriori* foi conduzida no G\*Power 3.1.9.7 para avaliar a sensibilidade do delineamento empregado. Considerando um tamanho de efeito médio ( $f = 0,25$ ),  $\alpha = .05$ , dois grupos entre sujeitos, quatro medidas repetidas, correlação moderada entre medidas ( $r = .50$ ) e  $\epsilon = 1$ , uma amostra total de 80 participantes resultou em poder estatístico elevado ( $1 - \beta = .99$ ), indicando sensibilidade adequada para detectar efeitos de magnitude média.

As respostas obtidas por meio do SAM foram analisadas separadamente para cada uma de suas três dimensões (*valência*, *arousal*, *dominância*), por representarem classes distintas de respostas avaliativas. Para cada dimensão, foram calculadas médias, desvios-padrão e intervalos de confiança das avaliações atribuídas às faces em função da raça e do gênero do estímulo, utilizando a raça

dos participantes como fator. Análises inferenciais avaliaram efeitos principais e interações entre fatores.

O desempenho no FAST foi analisado a partir da média das inclinações individuais das curvas de aquisição nos blocos consistente e inconsistente, calculada para cada grupo de participantes, permitindo comparar a fluência de aquisição de relações avaliativas sob contingências compatíveis ou incompatíveis com o viés racial socialmente dominante, em função da raça das crianças.

As respostas do Teste da Boneca foram analisadas por meio de estatísticas descritivas de frequência, seguidas de testes do *qui-quadrado* ( $\chi^2$ ) para examinar se os padrões de escolha diferiram significativamente entre as categorias de tom de pele das bonecas, conforme o contexto da pergunta.

## **Resultados**

### **SAM**

#### ***Valência***

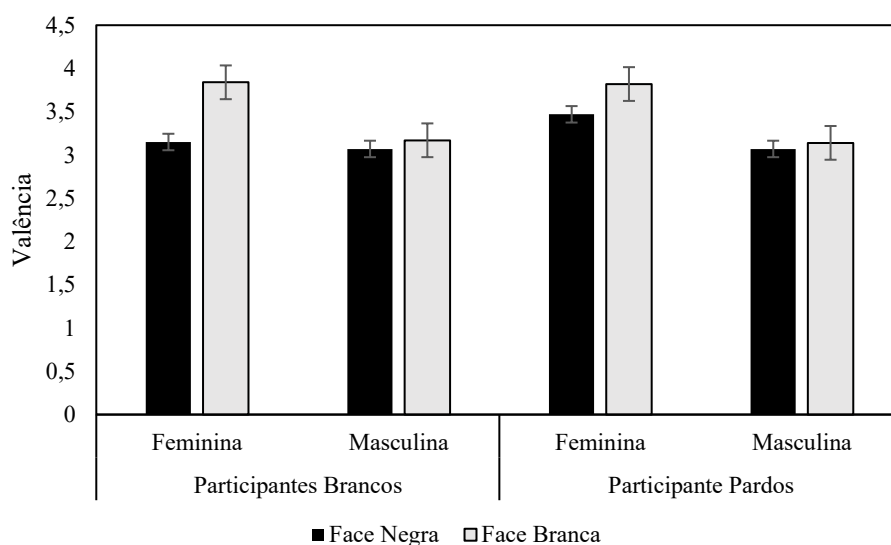
As avaliações de valência atribuídas às faces variaram sistematicamente em função da raça da face e do gênero da face. Entre as crianças brancas, as faces femininas brancas receberam as avaliações médias de valência mais elevadas ( $M = 3,84$ ;  $DP = 0,61$ ; IC95% [3,65, 4,02]), seguidas pelas faces femininas negras ( $M = 3,15$ ;  $DP = 0,58$ ; IC95% [2,97, 3,32]). As faces masculinas, tanto brancas ( $M = 3,17$ ;  $DP = 0,49$ ; IC95% [3,02, 3,32]) quanto negras ( $M = 3,07$ ;  $DP = 0,62$ ; IC95% [2,88, 3,26]), apresentaram médias mais baixas e bastante semelhantes entre si.

Entre as crianças pardas, observou-se um padrão descritivo semelhante. As faces femininas brancas novamente apresentaram as maiores médias de valência

( $M = 3,82$ ;  $DP = 0,59$ ;  $IC95\% [3,62, 4,02]$ ), seguidas pelas faces femininas negras ( $M = 3,47$ ;  $DP = 0,55$ ;  $IC95\% [3,29, 3,65]$ ). As faces masculinas foram avaliadas de forma menos positiva, tanto para faces brancas ( $M = 3,14$ ;  $DP = 0,40$ ;  $IC95\% [3,01, 3,27]$ ) quanto para faces negras ( $M = 3,01$ ;  $DP = 0,61$ ;  $IC95\% [2,81, 3,21]$ ). A Figura 3 apresenta as médias de valência atribuídas a faces negras e brancas, organizadas em função do gênero do estímulo facial e da raça dos participantes, evidenciando controles contextuais distintos sobre o responder avaliativo.

### Figura 3.

*Efeitos do gênero do estímulo facial e da raça da face sobre as avaliações de valência em crianças brancas e pardas.*



As avaliações de valência foram analisadas por meio de uma Análise de Variância Mista (ANOVA 2: Raça do Participante  $\times$  2: Raça da Face  $\times$  2 Gênero da Face), com raça da face (negra vs. branca) e gênero da face (feminino vs. masculino) como fatores intra-sujeitos, e raça do participante (branca vs. parda) como fator entre-sujeitos. A análise revelou um efeito principal da raça da face,

$F(1, 78) = 44,64, p < .001, \eta_g^2 = .075$ , e um efeito principal do gênero da face,  $F(1, 78) = 84,44, p < .001, \eta_g^2 = .152$ . Não foi observado efeito principal da raça do participante,  $F(1, 78) = 0,35, p = .554, \eta_g^2 = .002$ . Ademais, não foram identificadas interações significativas envolvendo a raça do participante. Em contraste, foi observada uma interação significativa entre raça da face e gênero da face,  $F(1, 78) = 15,61, p < .001, \eta_g^2 = .032$ , indicando que o efeito de uma dessas variáveis dependeu da outra.

As comparações pós-hoc, com correção de Bonferroni, foram conduzidas para descrever a direção dos efeitos principais. As comparações entre faces brancas e negras dentro de cada gênero indicaram que, entre faces femininas, as faces brancas receberam avaliações de valência significativamente mais elevadas do que as faces negras,  $t(79) = 7,28, p < .001$ . Em contraste, entre faces masculinas, não foi observada diferença significativa entre faces brancas e negras,  $t(79) = 1,66, p = .10$ . Adicionalmente, foram realizadas comparações entre faces femininas e masculinas dentro de cada categoria racial. Para faces brancas, as faces femininas foram avaliadas de forma significativamente mais positiva do que as faces masculinas,  $t(79) = 10,50, p < .001$ . De modo semelhante, para faces negras, as faces femininas também receberam avaliações de valência mais elevadas do que as faces masculinas, embora com menor magnitude do efeito,  $t(79) = 3,09, p = .003$ .

### ***Arousal***

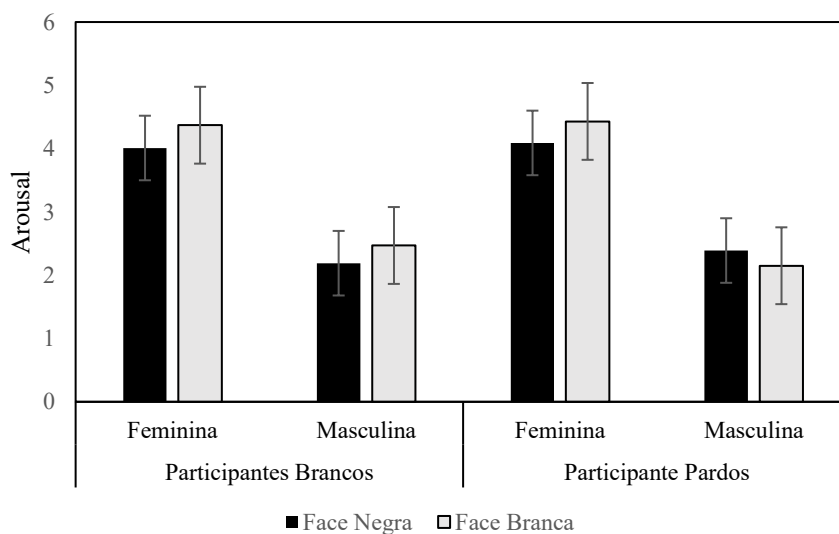
Entre participantes brancos, faces femininas receberam, de modo geral, escores mais elevados de *arousal* (indicando menor ativação) do que faces

masculinas, tanto para faces brancas (M = 4.37, DP = 0.39, IC95% [4.25, 4.49]) quanto para faces negras (M = 4.00, DP = 0.63, IC95% [3.81, 4.19]). Em contraste, faces masculinas apresentaram escores mais baixos, indicativos de maior ativação emocional, sendo essas médias semelhantes entre faces brancas (M = 2.47, DP = 0.34, IC95% [2.36, 2.58]) e faces negras (M = 2.19, DP = 0.47, IC95% [2.05, 2.34]).

Padrão semelhante foi observado entre participantes pardos. Faces femininas foram avaliadas com maiores escores de *arousal*, tanto para faces brancas (M = 4.43, DP = 0.40, IC95% [4.30, 4.57]) quanto para faces negras (M = 4.09, DP = 0.53, IC95% [3.91, 4.27]), indicando menor ativação emocional. Já faces masculinas receberam escores consistentemente mais baixos, refletindo maior ativação, tanto para faces brancas (M = 2.39, DP = 0.37, IC95% [2.27, 2.51]) quanto para faces negras (M = 2.15, DP = 0.46, IC95% [2.00, 2.31]). A Figura 4 apresenta as médias de *arousal* atribuídas a faces negras e brancas, organizadas em função do gênero do estímulo facial e da raça dos participantes, evidenciando controles contextuais distintos sobre a ativação eliciada pelos estímulos.

**Figura 4.**

*Efeitos do gênero do estímulo facial e da raça da face sobre as avaliações de arousal em crianças brancas e pardas.*



As avaliações de *arousal* foram analisadas por meio de uma ANOVA mista  $2 \times 2 \times 2$ , com raça do participante (branca, parda) como fator entre sujeitos, e raça da face (branca, negra) e gênero da face (feminina, masculina) como fatores intra-sujeitos. É importante destacar que, nessa dimensão da escala, escores mais baixos indicam maior ativação emocional (maior arousal), enquanto escores mais altos indicam estados mais calmos ou tranquilos.

A análise revelou um efeito principal significativo da raça da face,  $F(1, 78) = 33.21, p < .001, \eta_g^2 = .103$ , indicando que, de modo geral, faces negras foram associadas a escores mais baixos de arousal do que faces brancas, refletindo maior ativação emocional diante desses estímulos. Observou-se também um efeito principal robusto do gênero da face,  $F(1, 78) = 1454.90, p < .001, \eta_g^2 = .82$ , com faces masculinas recebendo escores significativamente mais baixos do que

faces femininas, indicando maior arousal associado a estímulos masculinos. Não foi observado efeito principal da raça do participante,  $F(1, 78) = 0.02$ ,  $p = .883$ ,  $\eta_g^2 < .001$ , nem interações significativas envolvendo raça do participante, raça da face e gênero da face. Esses resultados indicam que os padrões de responder afetivo observados foram consistentes entre crianças brancas e pardas.

As comparações pós-hoc, com correção de Bonferroni, foram conduzidas para descrever a direção dos efeitos principais. As comparações entre raças da face, realizadas separadamente para cada gênero, indicaram que faces negras evocaram maior ativação emocional (escores mais baixos) do que faces brancas, tanto para faces femininas,  $t(79) = 4.49$ ,  $p < .001$ , quanto para faces masculinas,  $t(79) = 4.15$ ,  $p < .001$ . Adicionalmente, as comparações entre gêneros da face, realizadas separadamente para cada raça, revelaram que faces masculinas foram avaliadas como mais ativadoras emocionalmente (escores mais baixos) do que faces femininas, tanto para faces brancas,  $t(79) = 36.6$ ,  $p < .001$ , quanto para faces negras,  $t(79) = 22.5$ ,  $p < .001$

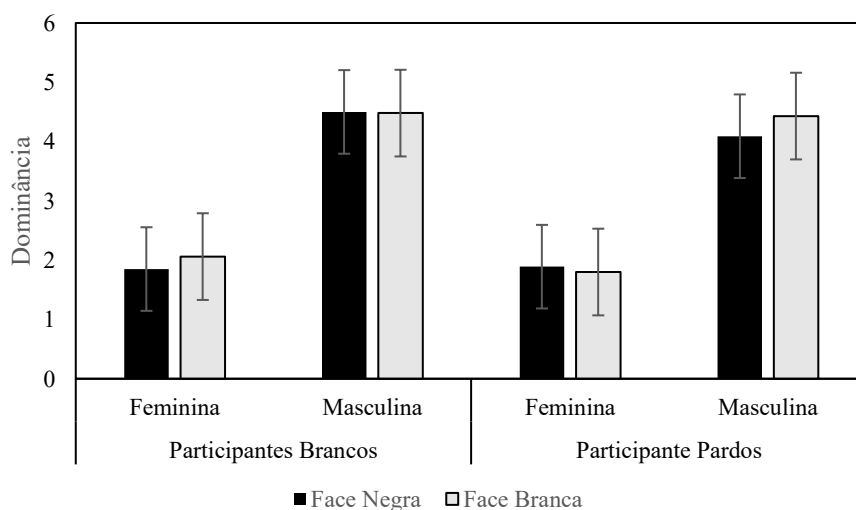
### ***Dominância***

De modo geral, faces masculinas foram avaliadas como mais dominantes do que faces femininas, independentemente da raça dos participantes e da raça da face. Entre participantes brancos, as médias de dominância atribuídas a faces masculinas foram elevadas tanto para faces brancas ( $M = 4.43$ ,  $DP = 0.27$ , IC95% [4.35, 4.51]) quanto para faces negras ( $M = 4.37$ ,  $DP = 0.27$ , IC95% [4.29, 4.45]). As avaliações de faces femininas nesse grupo foram substancialmente menores, com médias de 1.85 ( $DP = 0.43$ , IC95% [1.71, 1.98]) para faces negras e 2.06 ( $DP$

= 0.65, IC95% [1.86, 2.26]) para faces brancas. Padrão semelhante foi observado entre participantes pardos. As médias de dominância atribuídas a faces masculinas foram altas tanto para faces brancas ( $M = 4.50$ ,  $DP = 0.24$ , IC95% [4.42, 4.58]) quanto para faces negras ( $M = 4.48$ ,  $DP = 0.32$ , IC95% [4.37, 4.58]). Em contraste, as avaliações de faces femininas permaneceram baixas, com médias de 1.89 ( $DP = 0.38$ , IC95% [1.77, 2.02]) para faces brancas e 1.90 ( $DP = 0.59$ , IC95% [1.70, 2.10]) para faces negras. A Figura 5 apresenta as médias de *Dominância* atribuídas a faces negras e brancas, organizadas em função do gênero do estímulo facial e da raça dos participantes, evidenciando controles contextuais distintos sobre a percepção de dominância evocada pelos estímulos.

### Figura 5.

*Efeitos do gênero do estímulo facial e da raça da face sobre as avaliações de dominância em crianças brancas e pardas*



Os dados de dominância foram analisados por meio de uma ANOVA mista  $2 \times 2 \times 2$ , com raça do participante (Branca vs. Parda) como fator entre sujeitos, e raça da face (Branca vs. Negra) e gênero da face (Feminina vs. Masculina) como

fatores intra-sujeitos. A análise revelou um efeito principal robusto de gênero da face,  $F(1, 78) = 2272.42$ ,  $p < .001$ ,  $\eta_g^2 = .902$ , indicando que faces masculinas foram avaliadas como significativamente mais dominantes do que faces femininas, independentemente da raça do participante ou da raça da face. Não foram observados efeitos principais significativos da raça do participante,  $F(1, 78) = 0.09$ ,  $p = .763$ ,  $\eta_g^2 < .001$ , nem da raça da face,  $F(1, 78) = 0.68$ ,  $p = .412$ ,  $\eta_g^2 = .002$ . Também não foram encontradas interações estatisticamente significativas entre raça do participante e raça da face,  $F(1, 78) = 0.90$ ,  $p = .345$ ,  $\eta_g^2 = .002$ , entre raça do participante e gênero da face,  $F(1, 78) = 1.82$ ,  $p = .181$ ,  $\eta_g^2 = .007$ , ou entre raça da face e gênero da face,  $F(1, 78) = 3.05$ ,  $p = .085$ ,  $\eta_g^2 = .008$ . A interação de terceira ordem entre raça do participante, raça da face e gênero da face também não atingiu significância estatística,  $F(1, 78) = 2.03$ ,  $p = .158$ ,  $\eta_g^2 = .006$ .

Diante do efeito principal significativo de gênero da face identificado na ANOVA mista, foram conduzidas comparações post hoc pareadas (testes  $t$  pareados com correção de Bonferroni) para examinar as diferenças entre faces femininas e masculinas separadamente para cada raça da face. Os resultados indicaram que, para faces brancas, as avaliações de dominância foram significativamente mais altas para faces masculinas em comparação às faces femininas,  $t(79) = -44.10$ ,  $p < .001$ , evidenciando uma diferença robusta e consistente entre os gêneros. O mesmo padrão foi observado para faces negras, com faces masculinas sendo avaliadas como significativamente mais dominantes do que faces femininas,  $t(79) = -31.40$ ,  $p < .001$ .

## Teste da Boneca

As frequências de escolha no Teste da Boneca variaram de acordo com o contexto da pergunta. Na condição de preferência geral, a boneca negra foi escolhida com maior frequência ( $f = 53$ ) do que a boneca branca ( $f = 27$ ). Padrão semelhante foi observado na pergunta referente à escolha de amizade, com frequências próximas entre as categorias, embora ainda ligeiramente superiores para a boneca negra ( $f = 41$ ) em comparação à branca ( $f = 39$ ).

No contexto acadêmico avaliativo (atividade para nota), observou-se um padrão distinto, com maior frequência de escolha da boneca branca ( $f = 60$ ) em relação à boneca negra ( $f = 20$ ). Para a condição parte do grupo, a boneca negra foi novamente escolhida com maior frequência ( $f = 46$ ) do que a boneca branca ( $f = 34$ ). Por fim, na situação relacionada à Educação Física, a boneca negra apresentou a maior frequência de escolha ( $f = 57$ ), enquanto a boneca branca foi selecionada em menor número de respostas ( $f = 23$ ).

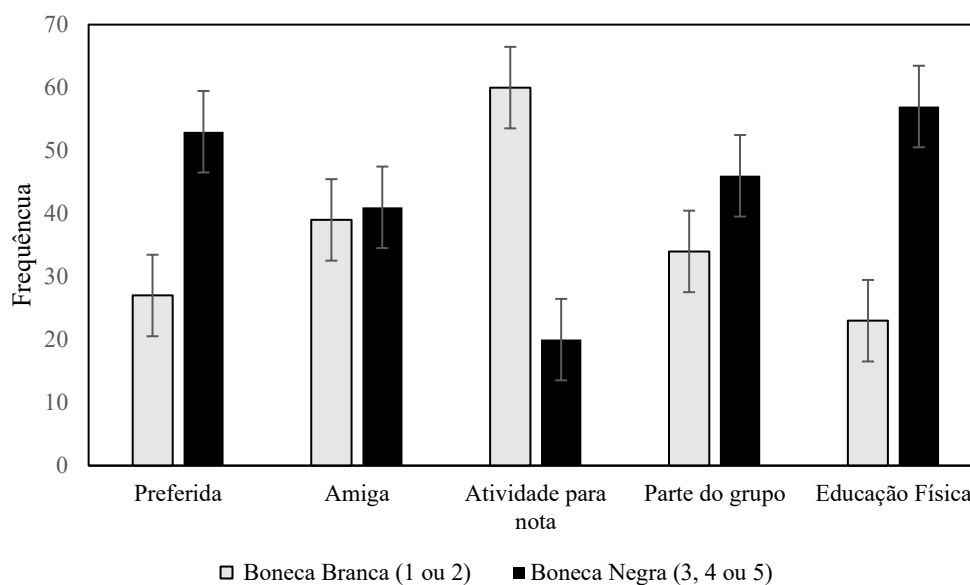
Para a pergunta sobre a boneca “Preferida”, observou-se uma diferença significativa na distribuição das escolhas, com maior frequência de seleção da boneca negra em comparação à boneca branca,  $\chi^2(1) = 15.63$ ,  $p < .001$ . De modo semelhante, na situação atividade para nota, a boneca branca foi escolhida com frequência significativamente maior do que a boneca negra,  $\chi^2(1) = 38.03$ ,  $p < .001$ . Também foi observada uma diferença significativa na situação Educação Física, com maior frequência de escolha da boneca negra,  $\chi^2(1) = 27.23$ ,  $p < .001$ .

Em contraste, para a pergunta “Escolheria alguma para ser sua amiga?”, não foi observada diferença significativa entre as frequências de escolha das bonecas branca e negra,  $\chi^2(1) = 0.03$ ,  $p = .874$ , indicando distribuição equivalente

das escolhas nessa situação. De forma semelhante, na situação “Escolheria alguma para fazer parte do grupo?”, a diferença entre as escolhas não atingiu significância estatística,  $\chi^2(1) = 3.03$ ,  $p = .08$ , embora tenha sido observada uma tendência em direção à maior escolha da boneca negra. A Figura 6 apresenta frequência que participantes selecionaram brancas (1 ou 2) e negras (3,4,5) para diferentes atividades.

### Figura 6.

*Frequência de escolhas da boneca branca e negra em função do contexto da tarefa no Teste da Boneca.*



### FAST

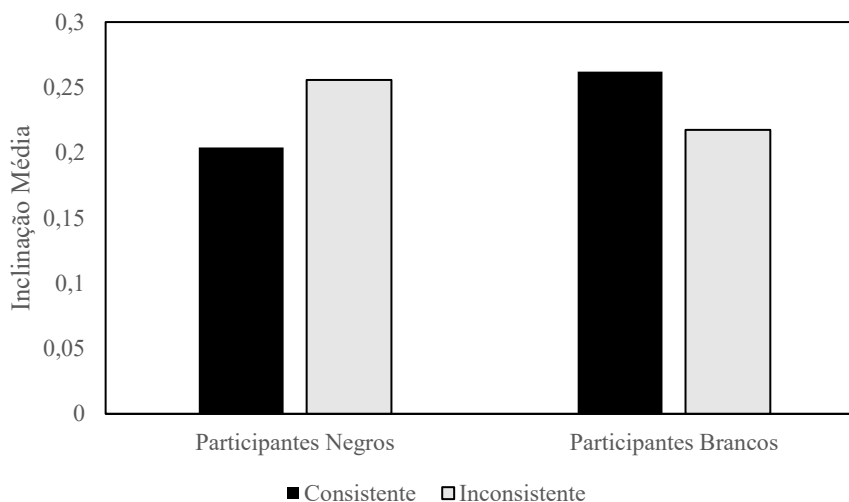
Entre participantes brancos, o desempenho médio foi superior no bloco consistente ( $M = .26$ ,  $DP = .07$ , IC95% [.24, .28]) em comparação ao bloco inconsistente ( $M = .30$ ,  $DP = .06$ , IC95% [.20, .24]), indicando aquisição relacional mais rápida quando as contingências estavam alinhadas a padrões relacionais culturalmente comuns, frequentemente descritos na literatura sobre viés racial.

Entre participantes pardos, o bloco inconsistente apresentou média mais elevada de desempenho ( $M = .26$ ,  $DP = .07$ , IC95% [.23, .28]) em relação ao bloco consistente ( $M = .20$ ,  $DP = .06$ , IC95% [.19, .22]), sugerindo maior velocidade de aquisição relacional sob condições inconsistentes para esse grupo.

A análise não revelou efeito principal da raça do participante,  $F(1, 78) = .75$ ,  $p = .39$ ,  $\eta_g^2 = .008$ , nem efeito principal do tipo de bloco,  $F(1, 78) = .47$ ,  $p = .49$ ,  $\eta_g^2 < .001$ . Esses resultados indicam que, quando considerados isoladamente, nem a raça do participante nem o tipo de bloco produziram diferenças sistemáticas no desempenho médio no FAST. Entretanto, foi observada uma interação robusta entre raça do participante e tipo de bloco,  $F(1, 78) = 72.617$ ,  $p < .001$ ,  $\eta_g^2 = .125$ . Esse resultado indica que o efeito do tipo de bloco sobre o desempenho no FAST variou em função da raça do participante, caracterizando um padrão cruzado de desempenho entre os grupos. A Figura 7 apresenta inclinação média nos blocos consistente e inconsistente no FAST, com base na raça dos participantes.

**Figura 7.**

*Inclinação Média nos Blocos Consistentes e Inconsistentes para participantes negros e brancos*



As comparações pós-hoc, com correção de Bonferroni indicaram que entre os participantes brancos, o desempenho no FAST foi significativamente superior no bloco consistente em comparação ao bloco inconsistente,  $t(42) = 5.65, p < .001$ . Esse resultado indica maior velocidade de aquisição sob condições consistentes para esse grupo. Em contraste, entre os participantes pardos, observou-se o padrão oposto, com desempenho significativamente superior no bloco inconsistente em relação ao bloco consistente,  $t(36) = -6.44, p < .001$ . Esse efeito inverso confirma que a direção do efeito do tipo de bloco variou sistematicamente em função da raça do participante.

### **Discussão**

O presente estudo investigou diferentes classes de respostas relacionadas a estímulos raciais em crianças, integrando medidas afetivas (Bradley & Lang, 1994), escolhas sociais explícitas (Clark & Clark, 1947) e desempenho relacional

implícito (Passarelli, Mizael et al., 2024). Os resultados indicaram que essas respostas não variaram de forma uniforme, mas estiveram sob controle de variáveis contextuais distintas, especialmente do gênero do estímulo e do tipo de tarefa, indicando que o preconceito racial não se manifesta como um padrão avaliativo único (Aboud, 1988, 1993; Bigler & Liben, 2007; Nesdale, 1999).

Observou-se que estímulos raciais exerceram controle diferencial sobre avaliações afetivas, padrões de ativação emocional e escolhas sociais, podendo se expressar tanto por desvalorização explícita quanto por formas topograficamente positivas de preferência, a depender do contexto funcional (Corrington et al., 2023; Katz et al., 1975). Esses achados sustentam uma compreensão do preconceito racial como um conjunto de respostas relacionais múltiplas, variáveis em topografia, função e contexto, cujos efeitos, ainda que heterogêneos, produzem consequências sistematicamente desvantajosas para determinados grupos (Guerin, 2003; Nogueira, 2007; Mizael & Sampaio, 2019; Saini & Vance, 2020).

Nas dimensões de valência, dominância e arousal, os resultados indicaram um padrão robusto de controle exercido pelo gênero da face, independentemente da raça dos participantes. Faces femininas foram avaliadas como mais positivas (ver Roche et al., 2025), menos dominantes (ver Cartwright et al., 2016) e menos ativadoras emocionalmente do que faces masculinas, enquanto faces masculinas evocaram avaliações mais negativas, maior dominância percebida e maior ativação emocional (Hugenberg & Bodenhausen, 2004; Bradley & Lang, 1994). Esse padrão foi consistente entre crianças brancas e pardas, sugerindo que o

controle exercido pelo gênero do estímulo facial constitui uma variável saliente na organização das respostas.

Sob uma leitura analítico-comportamental, esses resultados podem ser interpretados como reflexo de práticas culturais amplamente disseminadas que associam feminilidade a classes de estímulos relacionadas a cuidado, docilidade e menor ameaça, e masculinidade a classes relacionadas a força, autoridade e potencial aversivo (Bareket & Fiske, 2023; Glike & Fiske, 2001; Roche et al., 2025). Esse quadro ilustra as armadilhas e os riscos associados a topografias aparentemente positivas que, no longo prazo, acabam por legitimar iniquidades e sustentar relações de poder desproporcionais (Glick & Fiske, 2001; Phelan et al., 2010; Moya et al., 2007).

Resultados parcialmente convergentes foram observados entre as dimensões de valência e *arousal*, especialmente no que se refere às faces femininas. Em ambas as dimensões, faces femininas negras foram avaliadas de forma sistematicamente menos favorável do que faces femininas brancas, seja por meio de avaliações menos positivas de valência, seja por meio de maior ativação emocional (menores escores de *arousal*). Esse padrão sugere que, no caso de estímulos femininos, a raça da face exerceu controle consistente sobre diferentes classes de respostas avaliativas, indicando uma desvalorização relativa de faces femininas negras em comparação às brancas (Alexander & Carter, 2022; Mizael et al., 2021; Pacheco, 2008; Fanon, 2020).

Em contraste, para faces masculinas, as diferenças raciais não se expressaram de forma convergente entre as dimensões avaliativas, emergindo principalmente na ativação emocional, sem correspondência direta em

juízos de valência. Esse contraste entre faces femininas e masculinas sugere que o controle exercido por estímulos raciais é modulado pelo gênero, que opera como variável contextual na seleção de diferentes classes de respostas avaliativas. Esses padrões indicam que o preconceito racial não aparece como um efeito avaliativo global ou unitário, mas assume topografias distintas de responder avaliativo em função da combinação entre raça e gênero do estímulo. Assim, raça e gênero não devem ser tratados como categorias meramente descritivas, mas como variáveis contextuais que modulam o controle de estímulos sobre diferentes repertórios comportamentais, selecionando quais dimensões avaliativas tornam-se funcionalmente relevantes em cada contexto (Carbado et al., 2014; Crenshaw, 2013).

Quanto às escolhas sociais, mensuradas pelo Teste da Boneca adaptado, os resultados indicaram que os padrões de escolha variaram sistematicamente em função do contexto da atividade proposta. Em situações associadas a desempenho acadêmico avaliativo (atividade para nota), observou-se maior frequência de escolha da boneca branca, enquanto em contextos relacionados à Educação Física e preferência geral, a boneca negra foi escolhida com maior frequência. Em situações de amizade e escolha para grupo, não foram identificadas diferenças estatisticamente significativas entre as bonecas.

Esse padrão indica que as escolhas não foram aleatórias nem inconsistentes, mas sensíveis às contingências específicas evocadas por cada pergunta (Corrington et al., 2023; Auty & Lewis, 2004). Cada atividade operou como um contexto discriminativo distinto, evocando diferentes classes de respostas, como competência intelectual, habilidade física ou afinidade social.

Assim, a ausência de um viés racial uniforme não implica ausência de controle racial, mas evidencia que esse controle é contextual e funcionalmente específico (Nesdale et al., 2005; Bigler & Liben, 2007). Esse resultado é particularmente relevante porque sugere a organização precoce de estereótipos funcionais: estímulos associados à negritude foram mais frequentemente selecionados em contextos relacionados à força ou atividade física, enquanto estímulos associados à branquitude foram preferidos em contextos acadêmicos e avaliativos (Katz et al., 1975; Bar-Tal, 1996). Ainda que essas escolhas possam parecer “positivas” em determinados contextos, elas são potencialmente prejudiciais, pois restringem expectativas, oportunidades e trajetórias sociais (Clark & Clark, 1947; Nogueira, 2007; Blackson et al., 2022).

Os resultados do FAST revelaram um padrão funcionalmente relevante: enquanto crianças brancas apresentaram maior fluência relacional (taxa de aquisição) no bloco consistente com o viés racial socialmente dominante, crianças pardas apresentaram maior fluência no bloco inconsistente. Esse efeito cruzado foi confirmado por uma interação robusta entre raça do participante e tipo de bloco, replicando, em população infantil, a sensibilidade do FAST a histórias relacionais diferenciadas (Passarelli, Mizaël et al., 2024; Roche et al., 2025; Watters et al., 2023).). Esse achado dialoga diretamente com os resultados de Roche et al. (2025), nos quais participantes brancos apresentaram maior fluência no bloco consistente em comparação ao bloco inconsistente. No presente estudo, contudo, observa-se que, para crianças pardas, o padrão se inverte, com maior fluência na aquisição de relações que contradizem o viés racial culturalmente dominante.

Do ponto de vista funcional, esse padrão pode ser interpretado como indicativo de histórias relacionais distintas. Para o contexto dessa pesquisa, uma possibilidade é que crianças pardas estejam expostas a contingências culturais, educacionais ou familiares que favoreçam relações menos alinhadas ao viés dominante, seja por meio de experiências diretas, práticas de valorização identitária promovidas pela escola ou processos de contracontrole cultural (Mello & Fernandes, 2025). Em consonância com a literatura analítico-comportamental, esses resultados reforçam que o FAST é sensível não apenas à presença de avaliações negativas implícitas, mas também à organização diferencial dessas avaliações em função de variáveis contextuais, como a raça do respondente.

É fundamental cautela ao interpretar efeitos observados em medidas implícitas exclusivamente como indicadores de preconceito racial, especialmente quando tais efeitos assumem a forma de avaliações positivas direcionadas a membros do próprio grupo. Embora maior fluência relacional para estímulos racialmente categorizados como endogrupo possa refletir histórias de reforçamento e exposição cultural específicas, esse padrão, quando considerado isoladamente, não constitui evidência suficiente de preconceito racial nem permite inferir discriminação em contextos sociais concretos (Oswald et al., 2013; Fazio & Olson, 2003; Nosek, 2007). Do ponto de vista da Análise do Comportamento, o racismo envolve a organização de contingências culturais que produzem desvantagens sistemáticas, restringem o acesso a reforçadores sociais e mantêm tratamentos diferenciais persistentes em função da cor da pele (Mizael & Sampaio, 2019; Passarelli et al., 2023; Saini & Vance, 2020). Assim, um viés avaliativo implícito positivo em relação ao próprio grupo, como observado em crianças

negras ou pardas, não implica uma forma “reversa” de preconceito contra crianças brancas, na ausência de classes de respostas adicionais que resultem em exclusão, discriminação ou prejuízos funcionais. Avaliações positivas acerca de membros do endogrupo podem emergir precocemente no desenvolvimento, podendo (ou não) produzir efeitos sociais danosos (Passarelli, Mizael et al., 2024). Portanto, essas preferências se tornam relevantes para a análise do preconceito racial apenas quando a raça passa a exercer controle sobre respostas com consequências institucionais e desenvolvimentais, como preferências em ambiente de trabalho, respostas de discriminação, escolhas acadêmicas, atribuições de competência ou expectativas de desempenho.

Cabe destacar, contudo, que a própria noção de endogrupo pressupõe que o participante se reconhece como membro do grupo racial ao qual foi categorizado, o que não pode ser assumido sem verificação empírica. A formação da identidade racial em pessoas negras é um processo complexo, historicamente condicionado por relações de oposição ao grupo branco: a noção de “ser negro” pode estar sob controle não de redes relacionais vinculadas à negritude, mas da diferenciação em relação ao que é branco, de modo que atributos negativos historicamente associados ao grupo dominante se transferem por oposição para a identidade negra (Fanon, 2020; Gingles et al., 2022). Quando a identidade racial é constituída dessa forma, interpretar maior fluência relacional para faces negras como favorecimento do endogrupo pode ser equivocado. Limitação análoga aplica-se ao Teste da Boneca: a codificação binária das respostas de escolha pode mascarar efeitos de colorismo, uma vez que bonecas com tonalidades intermediárias e bonecas de pele mais escura são classificadas sob a mesma

categoria, podendo gerar a falsa impressão de viés positivo em relação a pessoas negras quando as escolhas favoráveis se concentram nas tonalidades fenotipicamente mais próximas do branco. Pesquisas futuras deveriam distinguir entre dois níveis de análise: quando o objetivo é examinar padrões de controle de estímulos sobre o responder avaliativo, a análise funcional dos dados é suficiente; quando o objetivo é interpretar esses padrões em termos de dinâmicas de endogrupo e exogrupo, torna-se necessário incluir avaliações do repertório de identificação racial dos participantes como variável independente.

Nesse sentido, uma definição de preconceito racial baseada exclusivamente em avaliações, explícitas ou implícitas, mostra-se necessariamente limitada, pois não captura outras classes de respostas igualmente centrais, como padrões de escolha, exclusão ou diferenciação funcional entre contextos (Corrington et al., 2023; Critchfield, 2011; Farmer et al., 2023; Slobodin et al., 2023). Esses limites reforçam a importância de abordagens multimodais, capazes de integrar medidas afetivas, implícitas e de escolha social, permitindo identificar quando e como a raça passa a operar como variável de controle indevida sobre o comportamento, bem como distinguir preferências contextuais de processos mais amplos de discriminação racial.

Os resultados do presente estudo contribuem para a compreensão do preconceito racial infantil ao evidenciar que diferentes facetas do responder racialmente relevante, avaliações afetivas, escolhas sociais e desempenho relacional implícito, estão sob controle de variáveis parcialmente distintas. A integração dessas medidas permitiu identificar padrões que não seriam detectados por uma única abordagem, revelando que o preconceito racial não se manifesta

como um repertório unitário, mas como um conjunto de respostas moduladas por contextos específicos, como gênero do estímulo, tipo de atividade e demandas da tarefa.

Do ponto de vista aplicado, esses achados indicam que intervenções voltadas à redução da discriminação racial devem ir além da modificação de avaliações explícitas e considerar repertórios relacionais implícitos e padrões de escolha funcionalmente relevantes. Por fim, ao demonstrar a utilidade de uma abordagem multimodal na infância, este estudo oferece subsídios empíricos para o desenvolvimento de estratégias educacionais e interventivas mais precisas, orientadas não apenas à identificação do preconceito racial, mas à modificação dos contextos nos quais a raça passa a exercer controle indevido sobre o comportamento.

-----

## Considerações Finais do Capítulo

### **Quando as medidas discordam: Interpretando o viés sob múltiplos controles**

O percurso empírico desenvolvido neste capítulo evidencia que o avanço na investigação do preconceito racial infantil não depende apenas do refinamento de instrumentos, mas do exame sistemático das condições sob as quais diferentes classes de respostas passam a ser selecionadas. Em continuidade direta ao estudo anterior, que demonstrou a estabilidade temporal do FAST e sua sensibilidade para identificar padrões relacionais implícitos em um subgrupo consistente de crianças, o último artigo deste capítulo amplia o foco da análise ao introduzir múltiplas medidas e variáveis contextuais no mesmo delineamento.

Enquanto o primeiro estudo respondeu a uma pergunta metodológica básica, mostrando que as medidas implícitas e explícitas são sensíveis e mantêm estabilidade suficiente para serem utilizadas em delineamentos longitudinais, o estudo multimodal desloca a análise para um nível funcional mais amplo. A questão deixa de ser apenas se há ou não viés racial e passa a ser sob quais condições o viés se manifesta, em que topografias de resposta e sob controle de quais variáveis contextuais.

Os resultados do estudo multimodal tornam explícito que avaliações afetivas, escolhas sociais e desempenho relacional implícito não variam de forma convergente. Pelo contrário, cada medida revelou padrões distintos de controle de estímulos. Avaliações na escala SAM foram fortemente moduladas pelo gênero do estímulo facial, com efeitos robustos e consistentes sobre valência, *arousal* e dominância, independentemente da raça das crianças. Esse padrão sugere que o gênero opera como uma variável contextual altamente saliente, capaz de organizar o responder avaliativo mesmo na presença de estímulos raciais.

Em contraste, a raça da face exerceu controle mais específico e dependente da dimensão avaliada. Para faces femininas, a raça modulou avaliações de valência e arousal de forma consistente, com faces negras sendo avaliadas de modo menos favorável do que faces brancas. Para faces masculinas, o controle racial foi mais evidente na dimensão de arousal, sem correspondência clara em valência. Esses resultados indicam que raça e gênero não operam como variáveis independentes, mas como elementos contextuais que se combinam para selecionar diferentes classes de respostas avaliativas.

As escolhas sociais, avaliadas por meio do Teste da Boneca modificado, reforçam essa interpretação contextual. As crianças não exibiram um padrão global de preferência racial, mas responderam diferencialmente em função da atividade proposta. Em contextos acadêmicos avaliativos, a boneca branca foi mais frequentemente escolhida; em atividades físicas e em preferências gerais, a boneca negra foi selecionada com maior frequência; em situações de amizade e pertencimento ao grupo, as escolhas se distribuíram de forma equivalente. Esses dados sugerem que o controle racial emerge associado a expectativas funcionais específicas, como competência intelectual ou habilidade física, e não como uma avaliação negativa generalizada.

O desempenho no FAST acrescenta uma dimensão adicional a essa análise. A interação robusta entre raça do participante e tipo de bloco evidencia que histórias relacionais distintas produzem padrões distintos de fluência na aquisição de respostas. Crianças brancas apresentaram maior fluência em relações consistentes com o viés racial socialmente dominante, enquanto crianças pardas exibiram maior fluência em relações inconsistentes. Esse efeito não pode ser adequadamente descrito como simples presença ou ausência de preconceito, mas como organização diferencial do responder relacional em função da posição social do respondente. Uma limitação a considerar é que o FAST utilizou

faces mistas em termos de gênero sem que essa variável fosse fatorada na análise, o que impede separar o controle exercido pela raça do controle exercido pelo gênero sobre o responder relacional. Dado que os resultados do SAM demonstraram que gênero e raça interagem na modulação do responder avaliativo, pesquisas futuras deveriam incluir o gênero das faces como fator no delineamento do FAST, permitindo identificar em que medida cada variável, isolada ou combinada, controla a fluência de aquisição relacional.

Esse achado dialoga diretamente com os resultados do artigo anterior, no qual apenas um subgrupo apresentou viés implícito consistente ao longo das reaplicações do FAST. No estudo multimodal, a introdução explícita da raça do participante como variável analítica revela que o FAST não é sensível apenas à direção do viés, mas à forma como diferentes histórias culturais e sociodemográficas modulam a aquisição de relações avaliativas. Assim, o FAST não funciona como um detector universal de preconceito, mas como um instrumento capaz de tornar visíveis diferenças funcionais entre repertórios relacionais.

Tomados em conjunto, os dois estudos indicam que a ausência de convergência entre medidas não deve ser interpretada como falha metodológica. Trata-se, antes, de uma propriedade do fenômeno investigado. O preconceito racial infantil não se apresenta como um repertório unitário, acessível por uma única classe de respostas. Ele emerge de forma fragmentada, distribuída entre avaliações afetivas, expectativas funcionais, escolhas sociais e padrões relacionais implícitos, cada qual sob controle de variáveis contextuais específicas.

Há ainda um aspecto metodológico que merece destaque e que atravessa silenciosamente todo o percurso empírico deste capítulo: a importância da flexibilidade analítica e da observação atenta do contexto durante a própria coleta de dados. O objetivo

inicial desta linha de pesquisa era investigar medidas de viés racial infantil, e a distribuição das faces por gênero foi concebida, a princípio, como um controle experimental necessário e suficiente. Não se tratava, naquele momento, de uma variável de interesse teórico central. No entanto, os efeitos sistemáticos observados durante a coleta, especialmente a modulação consistente do responder avaliativo pelo gênero do estímulo facial, só se tornaram evidentes porque o experimentador permaneceu sensível a padrões emergentes que não estavam previstos como hipóteses principais.

O mesmo processo ocorreu com a introdução da raça da criança como variável analítica. A separação entre crianças brancas e pardas não foi uma decisão tomada a priori, mas uma resposta analítica a diferenças funcionais observadas no desempenho relacional durante a coleta e a inspeção inicial dos dados. Sem essa sensibilidade ao controle contextual, os resultados teriam sido descritos como ausência de efeitos estatisticamente significativos, e o estudo poderia ser erroneamente interpretado como inconclusivo. Foi precisamente a atenção ao contexto de emissão das respostas, e não apenas aos instrumentos, que permitiu revelar padrões funcionais relevantes, como efeitos de endogrupo e histórias relacionais diferenciadas.

Esse ponto reforça uma implicação metodológica central: o momento da coleta não é um estágio meramente técnico do experimento, mas um componente analítico crítico. A investigação do preconceito racial, sobretudo na infância, exige que a pesquisadora observe ativamente como variáveis não previstas inicialmente passam a exercer controle sobre diferentes classes de respostas. Sem essa postura, delineamentos cuidadosamente planejados podem falhar não por falta de rigor, mas por insensibilidade às contingências efetivamente operantes. Assim, a produção de conhecimento relevante nesse domínio

depende tanto da precisão dos instrumentos quanto da capacidade do experimentador de ajustar o foco analítico diante dos efeitos que emergem no próprio curso da pesquisa.

Esses resultados recolocam questões que permanecem abertas. Em que momento do desenvolvimento variáveis sociodemográficas passam a exercer controle sistemático sobre determinadas classes de respostas e não sobre outras? Sob quais condições avaliações afetivas se tornam relevantes para escolhas sociais ou decisões institucionais? Quando padrões relacionais implícitos, identificados por medidas como o FAST, passam a se traduzir em práticas discriminatórias observáveis no cotidiano escolar? A essas questões soma-se uma variável sistematicamente negligenciada na pesquisa experimental sobre viés racial: as marcações sociais do próprio experimentador. Considerando que raça e gênero demonstraram exercer controle sobre o responder avaliativo dos participantes, é plausível que características fenotípicas e sociais de quem aplica os procedimentos também funcionem como variáveis contextuais que modulam o responder. Evidências nessa direção foram documentadas por Danso e Esses (2001), que observaram que participantes brancos com altos escores em dominação social apresentaram desempenho superior em tarefas cognitivas quando o aplicador era negro, padrão interpretado como responder sob controle da necessidade de afirmar superioridade diante de um membro do exogrupo. Pesquisas futuras deveriam identificar e reportar as características sociodemográficas da equipe de pesquisa e examinar sistematicamente sua influência sobre os resultados (Does et al., 2018; Pierre et al., 2023).

Do ponto de vista experimental, os dados deste capítulo sugerem que a investigação do preconceito racial exige delineamentos capazes de mapear relações funcionais entre múltiplas medidas, em vez de buscar convergência forçada entre elas. Do ponto de vista conceitual, indicam que compreender o preconceito racial implica analisar como diferentes

contextos selecionam diferentes repertórios, e não apenas identificar avaliações negativas isoladas.

À maneira proposta por Sidman, este capítulo não oferece respostas finais. Ele delimita com maior precisão o tipo de pergunta que pode ser feita a partir dos dados disponíveis e explicita os limites das inferências possíveis quando medidas são analisadas isoladamente. O capítulo seguinte parte dessas questões em aberto, investigando não apenas se o viés está presente, mas como ele se organiza funcionalmente quando diferentes medidas, contextos e histórias de aprendizagem são considerados simultaneamente.

####

## **Capítulo 4. Intervenções voltadas para Redução do Preconceito Racial em Crianças**

### **Background: A replicação sistemática**

Este capítulo marca um ponto de inflexão no desenvolvimento da presente tese. Não por estabelecer uma hierarquia de importância entre os estudos apresentados até aqui, mas por concentrar o objetivo que originalmente orientou o projeto submetido ao Programa de Pós-Graduação em Psicologia da UFSCar e que viabilizou a obtenção de financiamento junto à Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP). A proposta central consistia em investigar os efeitos de diferentes procedimentos de condicionamento avaliativo na redução do preconceito racial infantil, comparando estratégias baseadas em equivalência de estímulos e procedimentos com menor controle instrucional.

A compreensão desse objetivo exige retomar um problema empírico central da literatura sobre equivalência de estímulos envolvendo conteúdos socialmente carregados: o chamado paradigma das relações conflitantes. Esse paradigma descreve situações em que as relações treinadas entram em conflito com histórias pré-experimentais consolidadas, frequentemente resultando em dificuldades ou falhas na formação de classes de equivalência.

Estudos clássicos demonstraram que estímulos com funções previamente estabelecidas podem resistir à reorganização relacional quando inseridos em relações incompatíveis com essas funções. Trabalhos como os de Moxon et al. (1993) e Watt et al. (1991) mostraram que, em contextos de conflito relacional, a formação de classes de equivalência tende a não ocorrer, sugerindo limites funcionais importantes do procedimento quando confrontado com repertórios simbólicos historicamente estabelecidos.

No domínio de estímulos sociais, esse problema tornou-se particularmente relevante. Tentativas iniciais de aplicar o paradigma das relações conflitantes a conteúdos

racialmente carregados indicaram que a simples introdução de relações incompatíveis, como o treino de relações positivas envolvendo estímulos associados a avaliações sociais negativas, não era suficiente para produzir reorganização relacional. Nesses casos, observou-se a replicação dos efeitos descritos na literatura básica: ausência de formação de equivalência e manutenção das relações avaliativas previamente estabelecidas (Carvalho & de Rose, 2014; Haydu et al., 2015).

Um avanço empírico significativo ocorreu quando estudos subsequentes passaram a incorporar, de forma sistemática, parâmetros procedimentais derivados da literatura básica sobre equivalência de estímulos, reconhecidamente associados ao aumento da probabilidade de formação de classes. A adoção desses parâmetros no contexto de relações racialmente conflitantes permitiu demonstrar que, sob determinadas condições experimentais, classes de equivalência envolvendo estímulos racializados poderiam, de fato, ser formadas (Haydu et al., 2019; Mizael et al., 2016).

Nos estudos conduzidos por Mizael e colaboradores, a aplicação refinada do paradigma das relações conflitantes resultou na emergência de relações entre estímulos associados à categoria racial negra e estímulos de valência positiva, concomitantemente à eliminação de relações negro–negativo. Como consequência, observou-se redução consistente das diferenças de valência atribuídas a faces negras e brancas que, embora ainda presentes, deixaram de atingir significância estatística. Esses achados indicaram que o conflito relacional não constitui uma barreira absoluta à formação de equivalência, mas um desafio funcional sensível a parâmetros procedimentais específicos.

Assim, o paradigma das relações conflitantes passou a ocupar um lugar central na investigação do preconceito racial sob a perspectiva da equivalência de estímulos, deslocando a questão empírica da simples possibilidade de formação de classes para a

análise das condições sob as quais relações simbólicas socialmente consolidadas podem ser reorganizadas. É nesse contexto que se insere a presente tese, que amplia esse programa de pesquisa ao investigar, de forma sistemática, medidas, intervenções e padrões de generalização relacionados ao viés racial antinegro.

Antes, contudo, de avançar para a comparação entre diferentes formas de intervenção, impunha-se um passo metodológico incontornável: a replicação sistemática dos efeitos descritos por Mizael et al. (2016). Era necessário verificar não apenas a replicabilidade dos achados, mas também a estabilidade temporal das mudanças observadas e sua possível generalização para medidas distintas daquelas originalmente empregadas, incluindo o *Function Acquisition Speed Test* (FAST) e o Teste da Boneca.

Esse percurso teve início em um contexto atípico. No começo de 2021, quando ingressei no programa de doutorado, as atividades presenciais permaneciam suspensas em função da pandemia de COVID-19, o que exigiu que as primeiras tentativas de replicação fossem conduzidas de forma remota. A etapa inicial consistiu, portanto, em testar a viabilidade do procedimento em ambiente online. Uma chamada para participação em pesquisa resultou na adesão de apenas três voluntários, com idades entre 12 e 14 anos. Tratava-se de uma faixa etária mais familiarizada com recursos digitais e capaz de acessar o procedimento com menor necessidade de suporte, o que tornou esse grupo adequado para um estudo piloto.

Os responsáveis legais assinaram os termos de consentimento de forma remota, e os participantes forneceram assentimento. Essa etapa permitiu verificar que medidas explícitas poderiam ser aplicadas e reaplicadas online com estabilidade. Os participantes não apresentaram avaliações negativas em relação às faces negras, padrão que se manteve ao

longo de três exposições sucessivas, acompanhado de elevada consciência verbal acerca da inadequação moral da discriminação racial.

A análise da correspondência entre essas respostas explícitas e uma medida implícita, entretanto, foi inviabilizada por limitações técnicas. O FAST, configurado com restrição temporal, exigia acesso remoto ao computador do experimentador, e interrupções frequentes de conexão introduziram ruído suficiente para comprometer a interpretação dos dados. Ainda assim, essa etapa forneceu uma informação relevante para o programa de pesquisa: medidas explícitas podem ser implementadas remotamente, mantêm estabilidade ao longo de reaplicações e, sobretudo, não produzem mudanças quando avaliações negativas não estão presentes.

Com a retomada gradual das atividades presenciais ainda em 2021, tornou-se possível avançar para a replicação em ambiente escolar. Uma escola localizada em um município do interior de São Paulo aceitou participar da pesquisa, observando os protocolos sanitários vigentes. Foram disponibilizadas quatro salas, totalizando aproximadamente 120 alunos; entretanto, apenas 34 responsáveis autorizaram a participação de seus filhos. Diversos termos de consentimento devolvidos continham recusas explícitas, refletindo o contexto político e social sensível em torno de pesquisas sobre racismo naquele período.

As 34 crianças foram avaliadas em três ocasiões ao longo de uma semana. A partir dessas avaliações repetidas, foram selecionadas para a etapa interventiva as crianças que apresentaram viés racial antinegro de forma estável ao longo das reaplicações, obrigatoriamente presente no teste relacional. Nove crianças atenderam a esses critérios, definidos em função da própria natureza do procedimento interventivo, baseado na reorganização de relações derivadas.

O delineamento adotado seguiu uma lógica de medidas repetidas, com três avaliações pré-intervenção, três avaliações pós-intervenção e um follow-up realizado quinze dias depois. Os resultados desse estudo foram publicados em inglês na revista *Behavior and Social Issues*. A tradução do artigo e sua análise detalhada são apresentadas a seguir.

Esse estudo cumpriu, assim, uma dupla função: replicar efeitos previamente descritos na literatura e estabelecer as condições experimentais mínimas a partir das quais comparações entre diferentes procedimentos interventivos poderiam ser conduzidas de forma conceitual e metodologicamente responsável.

\*\*\*\*

## **Estudo 4 - Reduzindo o Viés Racial por meio do Paradigma de Relações**

### **Conflitantes: Uma Replicação Sistemática de Mizael et al. (2016)<sup>8</sup>**

Mizael et al. (2016) demonstraram que é possível reduzir avaliações negativas de faces negras e minimizar diferenças de valência entre avaliações de faces negras e brancas por meio do estabelecimento de relações de equivalência que entram em conflito com o viés racial. Replicamos esse procedimento em um delineamento de medidas repetidas e incluímos o Teste da boneca Adaptado e um teste implícito para avaliar generalização. Nove crianças com avaliações negativas de faces negras foram treinadas, ao longo de dois dias, a relacionar um símbolo positivo (A1) a um estímulo abstrato (B1) e B1 a imagens de faces negras (C1). Medidas de viés racial foram coletadas três vezes antes da intervenção e duas vezes após. Quinze dias após a última avaliação, realizou-se um seguimento com teste de manutenção das classes de equivalência e reavaliação das medidas de linha de base. Oito participantes formaram relações de equivalência entre faces negras e o símbolo positivo, mantidas duas semanas após a intervenção. Antes da intervenção, observou-se diferença significativa entre as avaliações de faces brancas e negras; após a intervenção, essa diferença deixou de ser significativa. Dois dos três participantes com viés implícito apresentaram redução do viés no pós-intervenção, e um dos quatro com viés no Teste da Boneca Adaptado mostrou generalização dos efeitos. No conjunto, os resultados replicam os achados de Mizael et al., embora a elevada disparidade entre os estímulos usados na triagem inicial e nos testes de generalização de mudança de função dificulte a interpretação. Pesquisas futuras são necessárias para consolidar procedimentos confiáveis de transformação de função e generalização.

---

<sup>8</sup> Passarelli, D. A., Roche, B., Mizael, T. M., Silvestre, M. H., & de Rose, J. C. (2024). Reducing racial bias using a conflicting relations paradigm: A systematic replication of Mizael et al. (2016). *Behavior and Social Issues*, 33(2), 737-764. <https://doi.org/10.1007/s42822-024-00181-6>

## Introdução

O preconceito racial é um problema social significativo, com impactos negativos profundos nos domínios econômico, social, psicológico e da saúde (Baker, 2020; Catagnus et al., 2022; Centers for Disease Control & Prevention [CDC], 2023; Paradies et al., 2015; Priest et al., 2021; Watson-Thompson et al., 2022). Dada a sua complexidade, definições claras são essenciais ao abordar esse fenômeno em pesquisas experimentais (Passarelli et al., 2023). Uma forma de definir o preconceito racial é como atitudes negativas direcionadas a um indivíduo ou grupo específico, culturalmente condicionadas e baseadas em atributos raciais (Mizael & de Rose, 2017; Nogueira, 2007).

O conceito de "atitudes" na literatura da análise do comportamento possui uma história complexa, sendo frequentemente tratado como um termo mentalista que requer definição funcional. Essa definição envolve compreender como um estímulo participa de uma rede relacional, levando à transferência ou transformação de funções avaliativas que variam ao longo de um espectro contínuo, e não necessariamente em uma polaridade discreta entre positivo e negativo (Mizael & de Rose, 2017; Roche et al., 2002). Essa formulação permite abarcar fenômenos como o colorismo e a interseccionalidade, nos quais o controle exercido por estímulos sociais não se expressa como avaliação binária, mas como padrões diferenciais de responder modulados pela combinação de múltiplas marcações sociais. Em um nível mais simples, e para os propósitos deste artigo, as atitudes podem ser entendidas como avaliações positivas ou negativas de

estímulos relevantes (Bohner & Dickel, 2011; Crano & Prislin, 2011; Mizael & de Rose, 2017; Schwarz, 2007).

As atitudes podem ser expressas por meio de respostas relacionais explícitas ou implícitas. As atitudes explícitas frequentemente refletem respostas afetivas dos indivíduos em interações sociais ou em questionários, geralmente produzidas sem pressão temporal, o que permite respostas relacionais complexas e elaboradas (Barnes-Holmes et al., 2010). Essas atitudes são comumente avaliadas por relatos verbais (por exemplo, afirmar explicitamente “não gosto de contratar pessoas negras porque são preguiçosas”) ou por questionários de atitudes (por exemplo, a Modern Racism Scale [MRS]; McConahay, 1986). Esse tipo de resposta também pode ser registrado por escalas do tipo Likert (por exemplo, Bradley & Lang, 1994) ou pela seleção de estímulos que simbolizam uma raça (por exemplo, imagens de faces negras, bonecas negras) na presença de estímulos positivos ou negativos (ver Mizael et al., 2016, 2020; ver também Clark & Clark, 1947). No estudo seminal de Clark e Clark (1947), por exemplo, a identificação racial e o preconceito foram examinados por meio do uso de bonecas para avaliar as respostas relacionais de crianças de seis anos a adjetivos positivos e negativos, revelando um viés pró-branquitude e indicando a presença de preconceito social desde cedo.

Por outro lado, respostas emitidas sob pressão temporal, moldadas principalmente por respostas verbais ou relacionais previamente reforçadas, podem ser consideradas atitudes implícitas. Em termos comportamentais, essas respostas implícitas são caracterizadas por respostas relacionais breves e imediatas, emitidas dentro de um intervalo de tempo restrito (ver Barnes-Holmes

et al., 2010; Watters et al., 2023). Essas respostas, controladas mais pela história de respostas relacionais do que por contingências atuais, podem ser mensuradas por procedimentos amplamente conhecidos como medidas de “testes implícitos” (por exemplo, o Implicit Relational Assessment Procedure [IRAP; Barnes-Holmes et al., 2010] e o Function Acquisition Speed Test [FAST; O’Reilly et al., 2012]). Esses testes buscam minimizar os efeitos confundidores da desejabilidade social na discriminação das avaliações de estímulos, ao alterar o equilíbrio de controle de contingências atuais para contingências passadas, embora ambas estejam sempre envolvidas em qualquer situação de resposta. Em particular, a desejabilidade social pode ser compreendida como responder a pistas atuais que funcionam como estímulos discriminativos sinalizando a disponibilidade de punição ou reforçamento para determinados tipos de respostas. Por exemplo, ao serem questionados se são racistas, indivíduos avaliam rapidamente pistas e normas sociais para fornecer uma resposta percebida como aceitável, visando evitar consequências negativas e, possivelmente, obter reforçamento positivo.

Partindo do pressuposto de que o preconceito racial consiste em atitudes negativas em relação a estímulos associados a uma determinada raça (Mizael & de Rose, 2017), uma forma de modificar essas atitudes é alterar a avaliação desses estímulos (Bohner & Dickel, 2011; Crano & Prislin, 2011). No campo da análise do comportamento, pesquisadores desenvolveram procedimentos para alterar funções afetivas de estímulos no contexto da mudança de atitudes sociais (por exemplo, Carvalho & de Rose, 2014; Mizael et al., 2016, 2021; Santos & de Rose, 2018, 2019). Em particular, esses estudos se baseiam nos princípios de equivalência de estímulos (Sidman, 1994) e no fenômeno da transferência ou

transformação de funções de estímulos (Dougher et al., 1994; de Rose et al., 1988). Isso é alcançado por meio do treino de discriminações condicionais (por exemplo, Sidman & Tailby, 1982), estabelecendo-se uma relação derivada comum entre estímulos com funções avaliativas conhecidas e estímulos cujas funções prévias a pesquisadora deseja modificar (ver Hughes et al., 2018; ver também Mizael et al., 2016, 2021).

Para fins ilustrativos, considere o seguinte procedimento paradigmático. Suponha que sejam estabelecidas duas classes arbitrárias de equivalência compostas por estímulos denominados A, B e C. Primeiro, o estímulo A é uma imagem que funciona como altamente apetitiva para muitos membros da população. Em seguida, o estímulo B é um estímulo abstrato neutro, relacionado a A durante o treino de discriminação condicional. Por fim, o estímulo C é uma imagem de uma face negra que já possui funções aversivas socialmente estabelecidas para alguns indivíduos e é relacionada a B. As exigências do treino de discriminação condicional podem levar à emergência de uma relação derivada entre os estímulos A e C, mediada pelo estímulo abstrato neutro, resultando na transferência de funções entre A e C (Dougher et al., 1994; Gandarela et al., 2020; de Rose et al., 1988). Em princípio, esse processo deveria resultar na redução das funções aversivas da imagem da face negra, o que pode ser mensurado por diversos métodos (Hughes et al., 2018; Mizael et al., 2016, 2021; Ribeiro et al., 2020; Santos & de Rose, 2018, 2019).

Como exemplo desse processo, Santos e de Rose (2018) conduziram um estudo com crianças em idade pré-escolar. As crianças foram treinadas a relacionar estímulos abstratos (C1 e C2) com personagens de desenho animado

de que gostavam (A1; por exemplo, Homem-Aranha) ou não gostavam (A2; por exemplo, Coringa). Essa relação foi estabelecida indiretamente por meio de um treino de discriminação condicional em que A1 e A2 foram ligados a formas geométricas (B1 e B2), que por sua vez foram relacionadas a figuras abstratas semelhantes a logotipos (C1 e C2). Posteriormente, as crianças foram expostas a dois recipientes rotulados com C1 e C2, ambos contendo o mesmo alimento. Os resultados mostraram preferência pelo alimento rotulado com o estímulo equivalente ao personagem favorito (C1), além de avaliações mais positivas do sabor desse alimento. Esse estudo exemplifica como a equivalência de estímulos pode influenciar respostas avaliativas.

Em um estudo conduzido por Carvalho e de Rose (2014) no contexto do viés racial, os pesquisadores buscaram modificar avaliações negativas de faces negras tornando-as equivalentes a estímulos com valência positiva. Quatro crianças que inicialmente avaliavam faces negras de forma negativa foram treinadas a relacionar indiretamente faces negras a estímulos positivos. Contudo, apenas uma criança formou uma classe de equivalência entre faces negras e estímulos positivos, demonstrando transferência de função ao avaliar as faces negras de forma mais positiva em uma escala Likert no pós-teste. Assim como em diversos estudos anteriores, o trabalho de Carvalho e de Rose buscou reverter a história prévia de condicionamento dos participantes utilizando um método posteriormente denominado por Mizael et al. (2016) de paradigma de relações conflitantes (Conflicting Relations Paradigm – CRP).

Estudos que utilizaram o CRP têm demonstrado a dificuldade de estabelecer relações de equivalência que se oponham à história pré-experimental

do indivíduo (por exemplo, Haydu et al., 2015; Moxon et al., 1993; Watt et al., 1991). Haydu et al. (2015), por exemplo, utilizaram o CRP para alterar as respostas afetivas de membros de torcidas organizadas de times de futebol brasileiros em relação a estímulos que representavam times rivais. O treino de discriminação condicional ao qual os participantes foram submetidos deveria resultar na emergência de relações de equivalência entre o emblema do time rival e palavras positivas, bem como entre o emblema do próprio time e palavras negativas. Contudo, apesar do uso de protocolos laboratoriais padrão, os pesquisadores não conseguiram estabelecer essas relações emergentes socialmente incongruentes (Strand & Arntzen, 2020). Na prática, o próprio efeito de transferência de funções, que tinha como objetivo reduzir atitudes antagonistas em relação ao grupo externo, não foi possível porque a formação de classes no laboratório foi impedida por relações de equivalência socialmente estabelecidas e incompatíveis envolvendo os mesmos estímulos.

Uma tentativa bem-sucedida de alterar a avaliação de faces negras utilizando o CRP foi relatada por Mizael et al. (2016; ver também Mizael et al., 2021). Treze crianças que inicialmente avaliavam faces negras de forma negativa em um pré-teste foram treinadas a relacionar um símbolo positivo (A1) a um estímulo abstrato (B1) e, subsequentemente, a relacionar o estímulo abstrato (B1) a faces negras (C1). Os autores utilizaram parâmetros que facilitam a formação de classes de equivalência e que normalmente não são empregados em estudos com CRP. Em particular, aplicaram treino misto das relações de linha de base após o treino de cada relação (Arntzen & Holth, 1997) e utilizaram o protocolo de treino “do simples ao complexo”, no qual testes de simetria são conduzidos

imediatamente após o treino de cada relação (Adams et al., 1993). Os resultados mostraram que todos os 13 participantes formaram classes de equivalência contendo imagens de faces negras e estímulos apetitivos. Além disso, após o treino, a diferença nas avaliações entre faces negras e brancas deixou de ser estatisticamente significativa no nível grupal. Esses achados foram replicados por Mizael et al. (2021).

A abordagem do CRP pode ser particularmente promissora quando utilizada com crianças, uma vez que respostas relacionais a estímulos raciais podem ainda não estar consolidadas em sua história de instruções ou reforçamento e, portanto, podem ser menos resistentes à mudança (Belisle & Dixon, 2020). No entanto, duas extensões desse corpo de trabalho parecem necessárias. Primeiro, é desejável uma avaliação da mudança de atitudes que apresente maior validade ecológica e de face. Segundo, considerando os avanços na área relacionados à avaliação da força de relações entre estímulos por métodos indiretos (por exemplo, FAST; O'Reilly et al., 2012), pode ser útil ampliar o conjunto de instrumentos utilizados para avaliar mudanças na força das respostas afetivas.

O presente estudo tem como objetivo replicar e estender os achados de Mizael et al. (2016, 2021), utilizando o Teste da Boneca Adaptado (Clark & Clark, 1947) como uma medida naturalística e com validade de face do viés racial. Além disso, foi utilizado um teste de estilo implícito como medida da força da relação entre estímulos raciais e estímulos de afeto negativo. Para esse fim, empregou-se o FAST, uma vez que medidas indiretas da força das relações entre estímulos demonstraram, em estudos laboratoriais, indexar a força das relações controladas

independentemente (Cummins et al., 2018, 2020). Esse instrumento tem demonstrado sensibilidade na avaliação de relações diretamente treinadas (O'Reilly et al., 2012) e de relações de equivalência derivadas (O'Reilly et al., 2013). Seu índice também distingue relações derivadas resultantes de diferentes quantidades de treino de discriminação condicional de linha de base (Cummins et al., 2018), bem como relações derivadas com diferentes distâncias nodais (Cummins & Roche, 2020).

## **Método**

### **Participantes**

Trinta e quatro estudantes do terceiro e quartos anos do ensino fundamental, com idades entre 8 e 9 anos, foram triados. Os critérios de inclusão incluíram a presença de viés racial em um teste relacional e em uma escala Likert (ver seção Ambiente, Equipamentos e Estímulos). Nove crianças com desenvolvimento típico (5 meninos;  $M = 8,22$  anos; intervalo: 8–9 anos), que com maior frequência relacionavam faces negras a um símbolo negativo e/ou faces brancas a um símbolo positivo no teste relacional, participaram do estudo.

Foi utilizado um delineamento de medidas repetidas, com seis medições repetidas por participante. A análise de poder estatístico, realizada no software G\*Power versão 3.1.9.4 (Faul et al., 2007), indicou um tamanho de efeito grande ( $f^2 = 0,46$ ), com nível de significância de 0,05 e poder estatístico de 95,13%. A Tabela 1 apresenta algumas características dos participantes, como idade, gênero, raça e anos de escolaridade, bem como a indicação das medidas nas quais os participantes apresentaram viés racial.

**Tabela 1.***Descrição dos Participantes e medidas que mostraram viés*

ID	Gênero	Idade	Identificação Racial (IBGE)	Ano Escolar	Presença de Viés Racial Antinegro			
					Teste Relacional	Escala Likert	Teste da Boneca	FAST
P1	Feminino	9 anos	Branco	4°	Sim	Não	Sim	Não
P2	Feminino	9 anos	Preto	4°	Sim	Sim	Sim	Não
P3	Feminino	8 anos	Pardo	3°	Sim	Sim	Não	Não
P4	Feminino	8 anos	Branco	3°	Sim	Sim	Não	Sim
P5	Masculino	8 anos	Pardo	3°	Sim	Sim	Não	No
P6	Masculino	8 anos	Branco	3°	Sim	Sim	Sim	Sim
P7	Masculino	8 anos	Branco	3°	Sim	Sim	Não	Sim
P8	Masculino	8 anos	Branco	3°	Sim	Sim	Não	Não
P9	Masculino	8 anos	Pardo	3°	Sim	Sim	Sim	Não

**Local, Equipamento e Estímulos**

Os dados foram coletados na biblioteca de uma escola municipal localizada em uma pequena cidade do estado de São Paulo, Brasil. Os estímulos e as consequências diferenciais foram apresentados por meio de um computador Dell G15, com processador Intel Core i5 operando a 2,50 GHz e equipado com 8 GB de memória. Os instrumentos utilizados foram os seguintes:

1. *Teste Relacional.* O teste foi administrado por meio do software MTS-III (Wallace, 2003), utilizando símbolos positivos e negativos (A1 e A2) como estímulos-modelo, faces negras (C1), um estímulo abstrato (C2) e faces brancas (C3). O teste consistiu em 16 tentativas, sem consequências diferenciais.
2. *Escala Likert.* As avaliações de valência de faces negras, faces brancas, estímulos abstratos, desenhos distratores, símbolos positivos e símbolos negativos foram coletadas por meio de uma escala Likert de 5 pontos,








utilizando emojis que expressavam emoções positivas e negativas. A apresentação da escala foi programada no software PsychoPy 3.0.

3. *Function Acquisition Speed Test (FAST)*. Esse instrumento avaliou a taxa de aprendizagem de categorização de estímulos em blocos de tentativas consistentes com as relações presumidas na história do participante e em blocos inconsistentes com essa história. Cada bloco consistiu em 50 tentativas, com instruções mínimas (ver O'Reilly et al., 2012). As tentativas envolveram a apresentação de um único estímulo no centro da tela: uma face negra, uma face branca, uma imagem com valência positiva ou uma imagem com valência negativa. Quatro exemplares de cada tipo de imagem foram apresentados aleatoriamente ao longo das tentativas (isto é, 16 estímulos no total), selecionados do International Affective Picture System (IAPS; Lang et al., 2005).
4. *Teste da Boneca Adaptado (Clark & Clark, 1947)*. Para o teste da boneca, foram utilizadas duas bonecas idênticas, com as mesmas roupas, mas com tons de pele diferentes, uma negra e outra branca. Foi desenvolvido um roteiro semiestruturado com sete perguntas utilizando adjetivos positivos (bom, legal, inteligente, bonito, forte, confiável e superior) e sete perguntas com adjetivos negativos contrastantes (mau, irritante, pouco inteligente, feio, fraco, perigoso e inferior).
5. *Intervenção*. As fases de treino e teste por pareamento ao modelo (MTS) foram as mesmas empregadas em estudos anteriores de Mizael et al. (2016, 2020). Os estímulos consistiram em A1 e A2 (símbolos positivo e negativo, respectivamente), B1 e B2 (estímulos abstratos) e classes de estímulos C1

e C2. A classe C1 consistiu em quatro faces negras (dois homens e duas mulheres), e C2 foi composta por um estímulo abstrato.

### Figura 1.

*Estímulos usados no Treino e nos Testes do Matching to Sample*

A1		A2	
			
B1		B2	
			
C1	C2	C3	
			

*Nota.* Faces foram apresentadas individualmente durante o treino.

### Procedimento

O procedimento foi aprovado pelo Comitê de Ética em Pesquisa da Universidade Federal de São Carlos (CAAE nº 48028721.0.0000.5504) e pela Secretária Estadual de Educação do Estado de São Paulo. A participação foi voluntária, com consentimento dos responsáveis legais e assentimento das crianças, sendo os estudos conduzidos mediante autorização da Secretaria Pública de Educação e da direção das escolas envolvidas. Foi utilizado um delineamento de medidas repetidas com fator intraindivíduos, que incluiu três avaliações de linha de base antes da intervenção (denominadas BL1, BL2 e BL3) e três avaliações adicionais após a intervenção (denominadas BL4, BL5 e seguimento). Cada conjunto de avaliações consistiu em: (1) um teste relacional,

utilizando um procedimento de pareamento ao modelo (*matching-to-sample* – MTS), para avaliar relações entre os conjuntos de estímulos A (símbolos) e C (faces e figura abstrata); (2) uma escala Likert para avaliar a preferência por faces negras e brancas; (3) o Function Acquisition Speed Test (FAST) para avaliar a força das relações entre faces negras e brancas e estímulos de valência positiva e negativa; e (4) o teste da boneca para avaliar preferência racial.

Os dados foram coletados por uma pesquisadora branca, de 29 anos. Cada avaliação de linha de base teve duração aproximada de 30 minutos, e o procedimento MTS durou cerca de 1 hora. As avaliações LB1, LB2 e LB3 foram realizadas ao longo de uma semana, na segunda-feira, quarta-feira e sexta-feira. A intervenção foi aplicada em dois dias consecutivos, com duração de 30 minutos por dia. Imediatamente após a intervenção, foi realizada a BL4, e dois dias depois foi conduzida a LB5. Uma avaliação de seguimento foi realizada 15 dias após a LB5, incluindo um teste de manutenção das classes de equivalência e a reaplicação das medidas de linha de base.

## **Medidas Avaliativas**

### ***Teste Relacional***

Os participantes foram selecionados por meio do procedimento de *matching-to-sample* (MTS), desenvolvido e utilizado por Mizael et al. (2016). Os participantes eram posicionados em frente a um computador. Um símbolo positivo ou negativo (isto é, uma imagem de “curtir” ou “não curtir”) era apresentado na tela, e os participantes eram instruídos a clicar nesse estímulo. Após o clique no estímulo-modelo, três estímulos de comparação apareciam abaixo: uma face negra (C1), um estímulo abstrato (C2) e uma face branca (C3), em posições

randomizadas. Em seguida, os participantes eram instruídos a clicar no estímulo de comparação que julgassem corresponder melhor ao estímulo-modelo.

O teste foi composto por um total de 16 tentativas, sendo oito com apresentação do estímulo-modelo A1 e oito com apresentação do estímulo-modelo A2, em ordem randomizada. O viés racial foi calculado por meio da seguinte fórmula:

$$v = [(B+) - (N+)] + [(N-) - (B-)],$$

em que  $v$  representa o índice de viés.  $B+$  e  $N+$  correspondem ao número de escolhas de faces brancas e negras, respectivamente, na presença do símbolo positivo, enquanto  $B-$  e  $N-$  correspondem ao número de escolhas de faces brancas e negras na presença do símbolo negativo. Um aumento no valor de  $v$  indica uma associação mais forte entre o símbolo positivo e faces brancas, bem como entre o símbolo negativo e faces negras (isto é, trata-se de um índice de preferência relativa).

Crianças que apresentaram um valor de  $v$  superior a 4 em três avaliações consecutivas de linha de base foram incluídas no estudo. A Figura 2 apresenta exemplos de dois tipos de tentativas do teste relacional.

## Figura 2.

*Exemplo dos dois tipos de tentativas apresentadas no Matching to Sample*



## EscaLa Likert

O estímulo a ser avaliado era apresentado no centro da tela. Abaixo, era exibida uma escala deslizante variando de um valor mínimo de 1 a um máximo de 5. Cada número da escala era acompanhado por um emoji que representava uma expressão emocional, sendo os valores 1 e 2 associados a emoções negativas, o valor 3 a emoção neutra, e os valores 4 e 5 a emoções positivas. Os participantes eram instruídos a atribuir uma pontuação a cada estímulo apresentado na tela utilizando a escala deslizante. Pontuações mais baixas indicavam valência negativa da imagem apresentada, enquanto pontuações mais altas indicavam valência positiva.

Os participantes avaliaram faces, símbolos, estímulos abstratos e estímulos distratores em ordem randomizada. Nesse teste, o viés racial foi considerado a partir do cálculo das médias das avaliações de valência para faces negras em comparação às faces brancas. Especificamente, o viés racial foi identificado quando os participantes atribuíram uma valência mais positiva às faces brancas

do que às faces negras. A Figura 3 apresenta o formato geral da escala Likert utilizada, em que cada tentativa exibia uma face branca, uma face negra (exemplificada na figura) ou um estímulo distrator.

**Figura 3.**

*Escala Likert utilizada para avaliação da valência de faces e dos distratores*



***Function Acquisition Speed Test (FAST; Cartwright et al., 2016)***

Os participantes iniciaram completando um bloco de treino que incluía estímulos não relacionados ao experimento principal. O teste consiste em uma tarefa diferencial de taxa de aprendizagem composta por dois blocos de aprendizagem: um no qual as contingências de reforçamento para o estabelecimento de uma classe funcional de estímulos são consistentes com a história social presumida do participante e outro no qual são inconsistentes com essa história. A ordem de apresentação dos blocos (consistente e inconsistente) foi contrabalanceada entre os participantes.

Em cada tentativa, um único estímulo era selecionado aleatoriamente pelo software e apresentado no centro da tela do computador. No bloco consistente, pressionar a tecla “z” na presença de faces brancas e imagens com valência positiva era reforçado. Da mesma forma, pressionar a tecla “m” na presença de faces negras e imagens com valência negativa também resultava em

reforçamento. No bloco inconsistente, as contingências de reforçamento eram invertidas, de modo que respostas de pressionar “z” eram reforçadas na presença de faces negras e imagens positivas, enquanto respostas de pressionar “m” eram reforçadas quando os participantes eram expostos a faces brancas e imagens negativas.

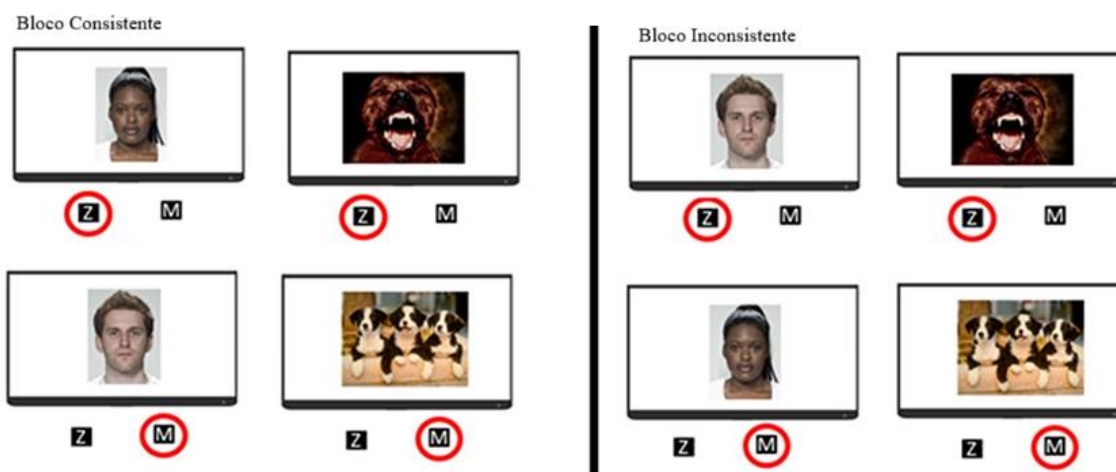
O software registrou tanto a acurácia quanto a latência das respostas com a finalidade de calcular uma inclinação de aprendizagem para cada bloco, com base no ângulo da linha de regressão da curva cumulativa de aprendizagem. Quando os participantes não respondiam dentro de 3.000 ms, a tentativa era encerrada e o feedback “incorreto” era apresentado na tela. A ausência de resposta foi registrada como uma resposta errônea para fins de cálculo do ângulo da linha de regressão da curva de aprendizagem. Feedback positivo ou negativo, na forma das palavras *correto* ou *errado* aparecendo na tela, foi apresentado para indicar o desempenho do participante. Após cada resposta, foi apresentado um intervalo de 500 ms, conforme apropriado, incluindo respostas omitidas (que geravam feedback negativo).

O efeito do FAST é geralmente concebido como o valor e a valência (positiva ou negativa) da inclinação da curva de aprendizagem do bloco consistente menos a inclinação da curva de aprendizagem do bloco inconsistente. Nesse caso, um escore de diferença positivo indica um viés pró-branco, enquanto um efeito FAST negativo indica um viés pró-negro. Para determinar a presença de viés racial, eliminando efeitos espúrios muito próximos de zero, foi estabelecido um limiar mínimo de efeito maior ou igual a +0,07 para a identificação de viés racial antinegro (conforme descrito por Marcelino, 2019). A Figura 3 apresenta

exemplos de tentativas administradas nos blocos consistente e inconsistente do FAST. As teclas “Z” e “M” reforçadas em cada bloco estão assinaladas por círculos vermelhos.

### Figura 3

*Exemplos de Tentativas apresentadas no Bloco Consistente e Inconsistente do FAST*



### Teste da Boneca Adaptado (baseado em Clark & Clark, 1947)

Foram utilizadas uma boneca branca e uma boneca negra como estímulos. A experimentadora sentava-se em frente ao participante, segurando uma boneca em cada mão, e fazia as seguintes perguntas: *Finja que essas bonecas são pessoas; aponte: Qual é branca/negra? Qual é boa/má? Qual é irritante/legal? Qual é inteligente/pouco inteligente? Qual é bonita/feia? Qual é confiável/perigosa? Qual é forte/fraca? Qual é melhor/pior? Qual se parece com você?*

A ordem das perguntas foi randomizada para evitar que os participantes respondessem sob a influência sistemática de qualidades positivas ou negativas.

Esses dados foram analisados qualitativamente. O viés racial foi considerado presente quando as crianças atribuíram com maior frequência atributos negativos à boneca negra ao longo das três avaliações de linha de base. Além disso, quando as crianças utilizaram a raça como justificativa ao serem questionadas sobre o motivo de suas escolhas, esse dado também foi registrado como indicador de viés racial.

A experimentadora registrou manualmente as respostas dos participantes no roteiro semiestruturado, indicando “negra” ou “branca” abaixo de cada pergunta, de acordo com a boneca selecionada. Posteriormente, essas respostas foram transferidas para uma planilha do Microsoft Excel para análise dos dados. Na planilha, cada participante foi representado por uma linha e cada adjetivo por uma coluna. Para cada adjetivo, registrou-se qual boneca foi selecionada e, ao final, foram somadas as frequências de escolha da boneca negra ou branca para adjetivos positivos e negativos.

Além disso, as crianças foram convidadas a justificar suas escolhas, por exemplo: “Por que ela é feia?”. A experimentadora registrou manualmente essas respostas no roteiro semiestruturado, anotando os casos em que características raciais foram utilizadas como justificativas para a atribuição de atributos positivos ou negativos.

### **Intervenção<sup>9</sup>**

Cada tentativa de MTS da intervenção envolveu a apresentação de um estímulo-modelo no centro superior da tela. Ao clicar no estímulo-modelo, os

---

<sup>9</sup> Esta tese pode ser lida de duas formas complementares: como um programa contínuo de pesquisa, em que cada estudo se articula com o anterior e os capítulos constroem

estímulos de comparação eram exibidos lado a lado abaixo dele. O estímulo-modelo permanecia na tela até que o participante selecionasse um dos estímulos de comparação. Respostas corretas eram seguidas pela apresentação de estrelas na tela, enquanto respostas incorretas resultavam em uma tela preta. O intervalo entre tentativas foi de 1,5 s, período durante o qual a tela permanecia preta. Foram administradas 16 tentativas para os treinos AB e BC, bem como para os testes BA, CB, AC e CA. A ordem das tentativas foi randomizada durante as fases de treino e de teste. Não houve consequências diferenciais nos blocos de teste; as respostas apenas iniciavam o intervalo entre tentativas.

Os participantes precisavam atingir pelo menos 15 respostas corretas em 16 tentativas para avançar para a etapa seguinte, com exceção do treino misto (AB, BC), que consistiu em 32 tentativas e cujo critério de aprendizagem foi estabelecido em 30 ou mais respostas corretas. As respostas produziram consequências diferenciais durante o treino, mas não durante os testes. Caso os participantes não atingissem o critério em qualquer fase de treino, o bloco de tentativas era repetido por até cinco vezes.

Durante os testes subsequentes de simetria (BA, CB), se o critério de 15 respostas corretas não fosse atingido, o bloco de treino correspondente era repetido até o critério, seguido por um novo teste. Esse ciclo de treino e teste foi

---

progressivamente o argumento empírico; ou como uma coletânea de artigos independentes, em que cada manuscrito pode ser consultado de forma isolada, sem prejuízo à compreensão. Em função dessa dupla possibilidade de leitura, todos os detalhes metodológicos foram preservados em cada estudo, de modo que o leitor que busca um artigo específico encontre nele as informações necessárias para compreender o procedimento sem recorrer a capítulos anteriores. O leitor que optar pela leitura integral da tese, no entanto, encontrará repetições procedimentais nos Estudos 5, 7 e 8, que compartilham o mesmo protocolo base de intervenção por *matching-to-sample*. Nesses casos, a repetição é deliberada e estrutural, não editorial.

repetido por até cinco vezes. Para os testes de transitividade e equivalência (AC, CA), apenas o teste era repetido, também por até cinco vezes, caso o critério não fosse alcançado.

Inicialmente, foi realizada uma sessão de pré-treino com estímulos não relacionados, com o objetivo de familiarizar os participantes com o procedimento. O procedimento principal iniciou-se com o treino AB, no qual um símbolo positivo (A1, sinal de positivo com o polegar) ou um símbolo negativo (A2, sinal de negativo com o polegar) era apresentado como estímulo-modelo, e os estímulos abstratos B1 e B2 eram utilizados como estímulos de comparação. A seleção de B1 na presença de A1 e de B2 na presença de A2 era reforçada. Durante o teste de simetria BA, B1 e B2 eram apresentados como estímulos-modelo, enquanto A1 e A2 funcionavam como estímulos de comparação.

Quando os participantes atingiam o critério de aprendizagem, iniciava-se o treino BC. Os estímulos-modelo eram B1 e B2, e os estímulos de comparação consistiam em um rosto negro (C1) e um estímulo abstrato (C2). A escolha de C1 na presença de B1 e de C2 na presença de B2 era reforçada, seguida pelo teste de simetria CB. Após o cumprimento do critério de aprendizagem, iniciava-se o treino misto ABBC. Esse treino consistiu em 32 tentativas, com 16 tentativas das relações AB e 16 das relações BC, apresentadas de forma intercalada. Os participantes avançavam para a fase de testes de equivalência ao alcançar 30 respostas corretas no treino misto (AB, BC).

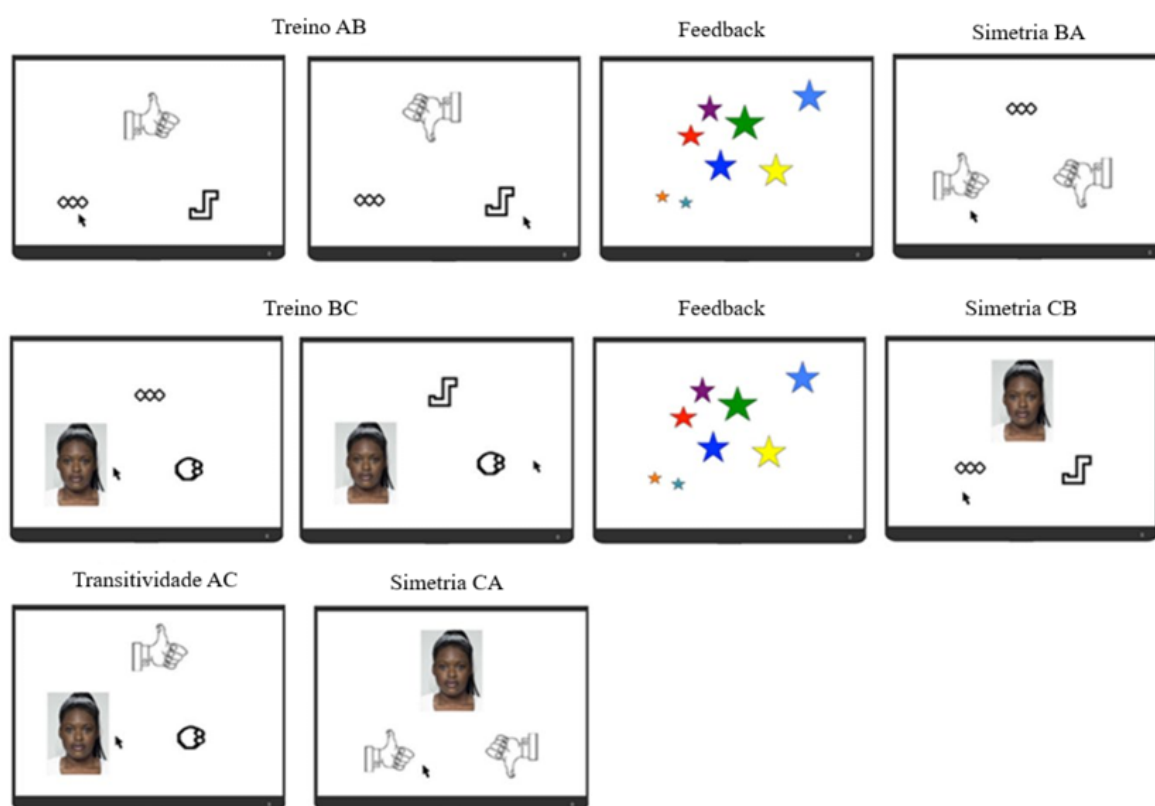
No teste de transitividade AC, os estímulos-modelo eram A1 ou A2, e os estímulos de comparação eram C1 e C2. Após o cumprimento do critério de aprendizagem, era conduzido o teste de equivalência CA, no qual C1 e C2

serviam como estímulos-modelo e A1 e A2 como estímulos de comparação. A formação de classes de equivalência foi determinada quando os participantes alcançavam pelo menos 15 respostas corretas nos testes AC e CA. A Figura 4 apresenta uma visão geral da intervenção, com alguns exemplos de tentativas.

Quinze dias após a intervenção, a manutenção das classes de equivalência foi avaliada por meio da reaplicação dos testes AC e CA. As quatro medidas de linha de base também foram reaplicadas.

#### Figura 4.

##### *Visão Geral da Intervenção*



#### Análise Inferencial

As análises estatísticas foram realizadas para cada medida (exceto o Teste da Boneca) utilizando o software IBM SPSS Statistics, versão 29.0.0.0 (241). O

teste de Friedman para amostras relacionadas (two-way) foi selecionado como uma alternativa não paramétrica robusta para a análise de medidas repetidas, considerando tanto as limitações do tamanho amostral quanto as violações do pressuposto de normalidade observadas em alguns conjuntos de dados. O coeficiente de Kendall (W) foi utilizado para estimar o tamanho do efeito, fornecendo a magnitude das diferenças observadas. Adicionalmente, para controlar o erro associado a múltiplas comparações, foram realizados contrastes de Wilcoxon com correção de Bonferroni.

## **Resultados**

### **Intervenção**

Oito participantes (88,88%) formaram com sucesso classes de equivalência, alcançando mais de 15 respostas corretas nos testes AC e CA (o participante P9 não formou classes). Na avaliação de seguimento realizada 15 dias após a intervenção, todos os oito participantes demonstraram a manutenção das classes de equivalência de acordo com os critérios estabelecidos. Assim, considera-se que a intervenção foi implementada com sucesso em oito dos nove participantes. A partir disso, passa-se a examinar o desempenho dos participantes nas diferentes medidas de viés racial ao longo dos múltiplos pontos temporais, do pré ao pós-intervenção.

### **Teste Relacional**

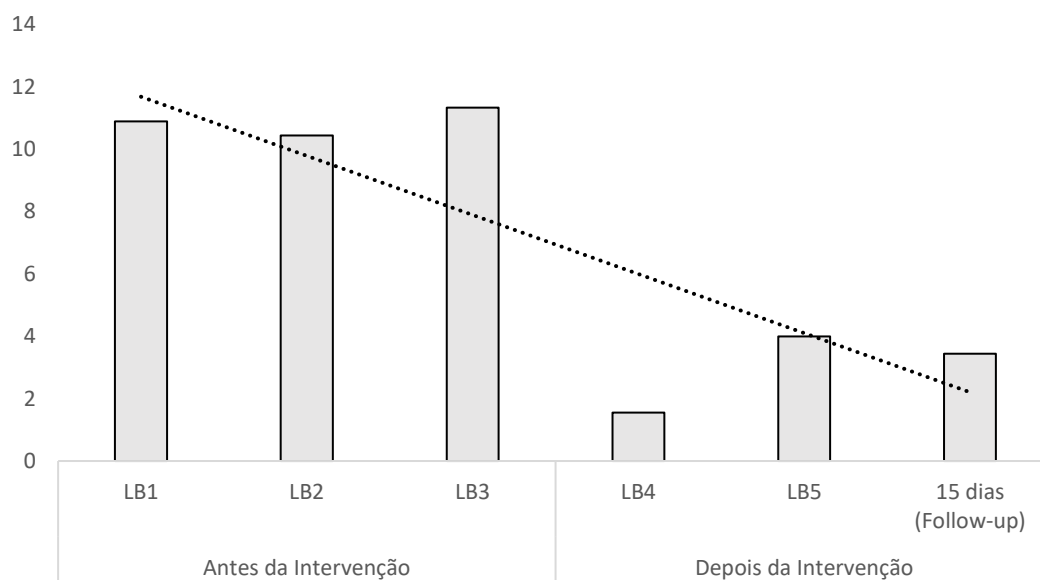
Antes da intervenção (LB1, LB2, LB3), o desempenho do grupo apresentou um índice médio de viés racial de 10.89 ( $DP = 2.57$ ; IC 95% [9.48, 12.52]). Nas linhas de base pós-intervenção, o índice médio de viés racial reduziu-se para 2.78

( $DP = 4.91$ ; IC 95% [-2.28, 5.94]). No seguimento realizado 15 dias após a intervenção, esse índice apresentou média de 3.44 ( $DP = 5.74$ ; IC 95% [-0.11, 6.89]). A Figura 5 ilustra o índice médio de viés racial obtido pelos participantes nas medidas pré- e pós-intervenção, bem como no seguimento. A linha de tendência indica uma redução do índice de viés racial após a intervenção.

A análise de Friedman revelou diferenças significativas entre as condições de linha de base,  $\chi^2 (5, N = 9) = 32.941, p < .001$ . O coeficiente de Kendall ( $W$ ) indicou um tamanho de efeito elevado ( $W = 0.73$ ), sugerindo variações substanciais no índice de viés ao longo dos diferentes pontos temporais. As comparações pareadas com correção não indicaram diferenças significativas entre as medidas de linha de base pré-intervenção (LB1, LB2, LB3). Diferenças estatisticamente significativas foram observadas entre LB1 e LB4 ( $z = 3.52, p = .06$ ), LB1 e LB5 ( $z = 3.09, p = .03$ ), LB2 e LB4 ( $z = 3.09, p = .016$ ), LB3 e LB4 ( $z = 3.91, p = .001$ ), LB3 e LB5 ( $z = 3.47, p = .08$ ) e entre LB3 e o seguimento ( $z = 3.15, p = .025$ ). Para os contrastes entre as linhas de base pós-intervenção (LB4, LB5 e LB6), não foram encontradas diferenças significativas.

**Figura 5.**

*Índice de Viés Racial Antinegro nas Linhas de Base Antes e Depois da Intervenção*

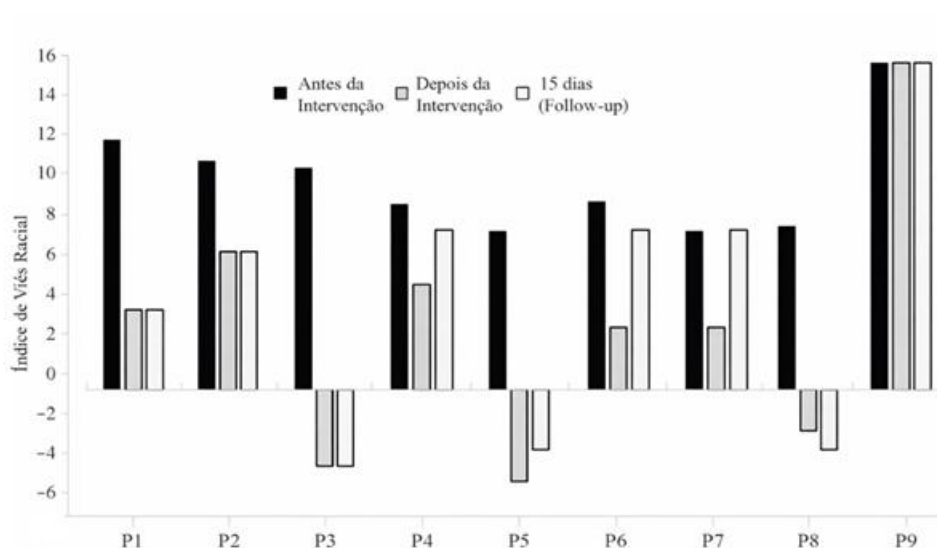


A análise individual é apresentada na Figura 6, que mostra, para cada participante, as médias dos valores de  $b$  no pré- e pós-intervenção. Os participantes P1 a P8 apresentaram valores de  $b$  pré-intervenção variando entre 8 e 12. Após a intervenção, os escores diminuíram para todos esses participantes, com a magnitude da redução ( $b_{\text{pré}} - b_{\text{pós}}$ ) variando de 4 a 15 pontos. Destaca-se que, para todos esses participantes, a frequência de escolhas da face negra quando o estímulo-modelo era o símbolo negativo caiu para zero nas avaliações pós-intervenção. Nesse momento, eles tenderam a escolher o estímulo abstrato na maioria desses ensaios. Três participantes passaram a escolher a face negra na maioria dos ensaios com o estímulo-modelo positivo, enquanto seis participantes continuaram a escolher a face branca na maioria desses ensaios, resultando em uma menor redução do valor de  $b$ . Cinco participantes mantiveram o valor reduzido

de  $b$  no seguimento, enquanto, para outros três participantes (P4, P6 e P7), o valor de  $b$  retornou a níveis próximos aos observados no pré-intervenção. No entanto, mesmo para os participantes em que o valor de  $b$  aumentou, não houve mais escolha da face negra quando o símbolo negativo foi apresentado como estímulo-modelo.

### Figura 6

*Viés Racial Médio Antes da Intervenção, depois da Intervenção e no seguimento de 15 dias após a intervenção (Follow-up)*



Nota. O rótulo “Antes da Intervenção” refere-se ao desempenho médio nas três avaliações pré-intervenção, e “Depois da Intervenção” refere-se à média das duas avaliações pós-intervenção. A avaliação de follow-up foi realizada 15 dias depois.

### Escala *Likert* (Valência)

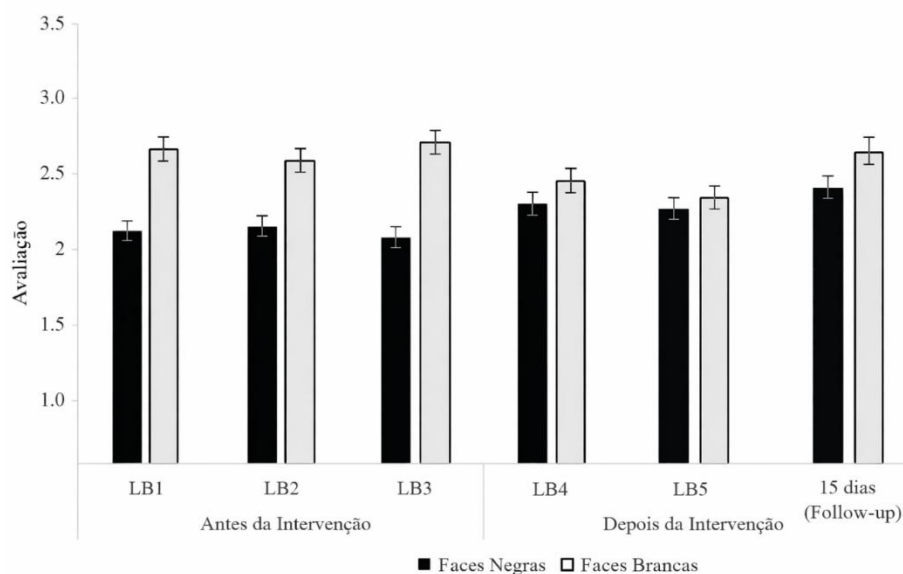
Os participantes tenderam a avaliar negativamente tanto as faces brancas ( $M = 2.84$ ,  $DP = 0.44$ , IC 95% [2.55, 3.14]) quanto as faces negras ( $M = 2.15$ ,  $DP = 0.48$ , IC 95% [1.89, 2.42]) nas três linhas de base pré-intervenção. De modo geral, as faces negras receberam avaliações mais baixas do que as faces brancas. É

interessante notar que os estímulos distratores receberam avaliações mais positivas em comparação às faces, apresentando médias entre 3.8 e 4.2 nas linhas de base antes e após a intervenção.

Nas duas linhas de base pós-intervenção, houve uma redução nas diferenças entre as avaliações das faces, com as faces negras sendo avaliadas de forma ligeiramente mais positiva ( $M = 2.41$ ,  $DP = 0.56$ , IC 95% [2.04, 2.82]) e as faces brancas sendo avaliadas de forma ligeiramente mais negativa ( $M = 2.55$ ,  $DP = 0.42$ , IC 95% [2.14, 2.99]). No seguimento, as faces brancas apresentaram um leve aumento nas avaliações ( $M = 2.69$ ,  $DP = 0.61$ , IC 95% [2.33, 3.05]), enquanto as faces negras mantiveram avaliações semelhantes às do pós-intervenção ( $M = 2.44$ ,  $DP = 0.69$ , IC 95% [2.00, 2.89]). A Figura 7 indica a valência das faces negras e brancas ao longo das linhas de base antes e após a intervenção.

### Figura 7

*Valência das faces negras e brancas antes e depois da intervenção*



Foi realizado um teste de Friedman para amostras relacionadas a fim de avaliar as diferenças nas avaliações das faces negras e brancas ao longo das condições LB1 a LB6. Os resultados revelaram uma diferença estatisticamente significativa nas avaliações das faces ao longo das linhas de base,  $\chi^2(11, N = 9) = 38.03$ ,  $p < .001$ . O Coeficiente de Kendall (W) indicou um tamanho de efeito médio de 0.38, sugerindo variações substanciais nas avaliações das faces ao longo dos diferentes pontos temporais. As comparações pareadas indicaram diferenças significativas entre faces negras e brancas em LB1 ( $z = -2.96$ ,  $p = .03$ ), LB2 ( $z = -2.41$ ,  $p = .014$ ) e LB3 ( $z = -3.33$ ,  $p < .001$ ). No entanto, após a correção de Bonferroni, diferenças significativas foram observadas apenas no contraste entre faces negras e brancas em LB3 ( $p = .037$ ).

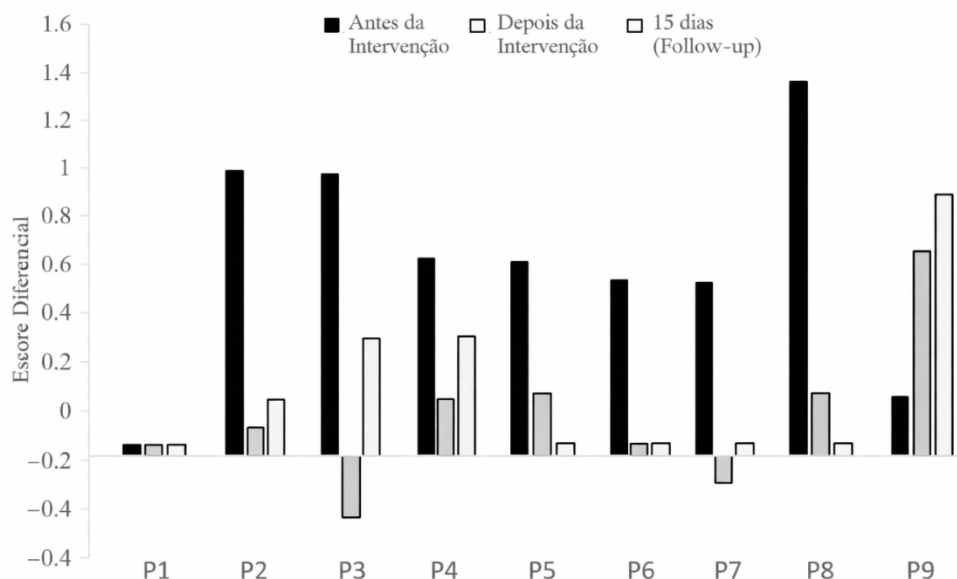
Oito participantes apresentaram viés racial em suas avaliações na escala Likert, atribuindo escores mais baixos às faces negras do que às faces brancas nas três avaliações pré-intervenção (LB1, LB2, LB3). A única exceção foi o participante P1, que atribuiu a mesma avaliação para ambas as faces nas linhas de base pré e pós-intervenção. A Figura 7 apresenta as médias das diferenças entre as avaliações de faces negras e brancas; os valores foram obtidos subtraindo-se, para cada participante, as avaliações das faces negras das avaliações das faces brancas no pré-intervenção, pós-intervenção e no seguimento. Diferenças positivas indicam que as faces brancas foram avaliadas mais positivamente do que as faces negras.

A Figura 7 mostra uma redução acentuada na diferença entre as avaliações pós-intervenção e pré-intervenção para todos os participantes que formaram classes de equivalência. O participante P9, que não formou classes de equivalência, foi uma exceção, apresentando aumento na diferença das avaliações. Com grande

cautela, essa observação pode ser tratada como uma observação controlada da falha previsível da intervenção. Entre os participantes que formaram classes de equivalência, as diferenças diminuíram para valores inferiores a 0,4, aproximando-se de zero para a maioria dos participantes. No seguimento, dois desses participantes apresentaram um pequeno a moderado aumento nas diferenças, ainda assim permanecendo abaixo dos níveis observados no pré-intervenção.

### Figura 7

*Diferença média nas avaliações em escala Likert entre faces negras e faces brancas*



*Nota.* O eixo Y apresenta o escore diferencial de avaliação, isto é, a diferença de valência atribuída às faces brancas e às faces negras. O rótulo “Antes da Intervenção” refere-se ao desempenho médio nas três medições de linha de base antes da intervenção, e “Depois da Intervenção” refere-se à média das duas medições de linha de base após a intervenção. A avaliação de follow-up foi realizada 15 dias depois. Escores diferenciais positivos representam viés antinegro. Escores negativos representam viés pró-negro.

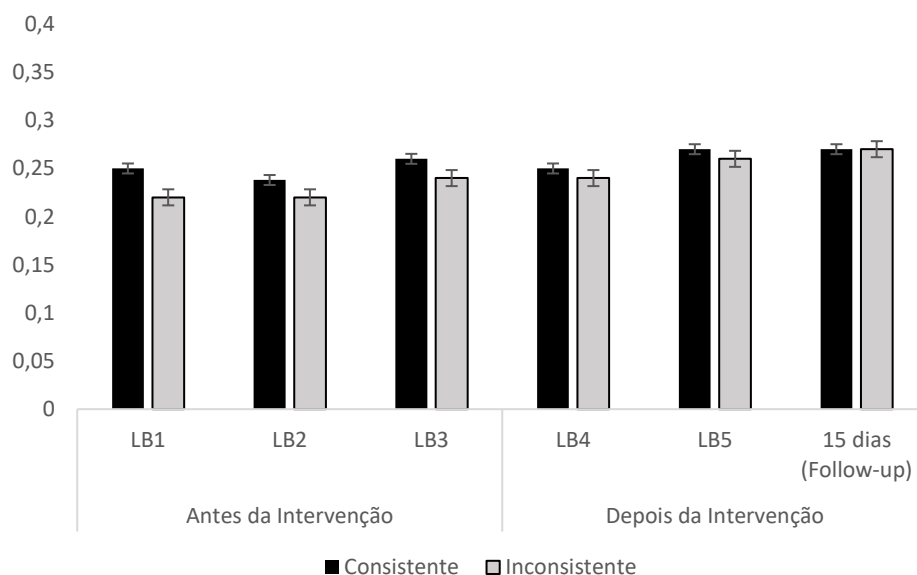
### **Function Acquisition Speed Test (FAST)**

No nível grupal, o desempenho nos blocos consistente e inconsistente do FAST é apresentado na Figura 8. De modo geral, observou-se uma tendência de

os participantes apresentarem desempenho ligeiramente mais fluente no bloco consistente ( $M = 0.26$ ,  $DP = 0.04$ , IC 95% [0.23, 0.29]) em comparação ao bloco inconsistente ( $M = 0.24$ ,  $DP = 0.04$ , IC 95% [0.21, 0.26]) ao longo das três iterações pré-intervenção, indicando um viés antinegro consistente, porém muito pequeno.

### Figura 8

*Desempenho nos Blocos Consistentes e Inconsistentes do FAST Antes e Após a Intervenção*



Nas duas linhas de base pós-intervenção, a fluência no bloco consistente ( $M = 0.26$ ,  $DP = 0.04$ , IC 95% [0.23, 0.28]) permaneceu ligeiramente superior à do bloco inconsistente ( $M = 0.25$ ,  $DP = 0.06$ , IC 95% [0.21, 0.29]). Durante o follow-up, a fluência média foi idêntica nos blocos consistente ( $M = 0.28$ ,  $DP = 0.06$ , IC 95% [0.24, 0.31]) e inconsistente ( $M = 0.28$ ,  $DP = 0.08$ , IC 95% [0.23, 0.33]).

Um teste de Friedman para amostras relacionadas indicou ausência de diferenças significativas entre os blocos ao longo das linhas de base,  $\chi^2$  (11,  $N = 9$ )

= 6,16,  $p = .86$ , com tamanho de efeito muito pequeno ( $W = 0.06$ ). As comparações pareadas confirmaram a inexistência de diferenças significativas nos contrastes entre blocos e linhas de base (LB). Assim, apesar das pequenas diferenças observadas, estas não foram suficientes para alcançar significância estatística.

Apenas três participantes (P4, P6 e P7) apresentaram viés antinegro estável no desempenho no FAST durante as três avaliações pré-intervenção. P4 e P6 exibiram um viés antinegro substancial, com efeito FAST superior a 0.10 nas três avaliações pré-intervenção. Após a intervenção, P4 apresentou um leve viés pró-negro, com efeito FAST de -0.02 em LB4. Contudo, em LB5, o efeito FAST foi de 0.02 e, durante o follow-up, aumentou para 0.05, indicando um retorno a um leve viés pró-branco. Por outro lado, P6 apresentou viés pró-negro em LB4 (*efeito FAST* = -0.03), LB5 (*efeito FAST* = -0.04) e durante o follow-up (*efeito FAST* = -0.02).

O participante P7 apresentou efeito FAST variando entre 0.05 e 0.08 nas três avaliações das linhas de base pré-intervenção ( $M = 0.06$ ). Nas duas linhas de base pós-intervenção, houve uma redução do efeito FAST, que diminuiu para 0.03 e 0.00, respectivamente. Entretanto, durante a avaliação de follow-up, o efeito FAST aumentou para 0.07, indicando um retorno ao padrão pré-intervenção.

### **Teste da Boneca**

O Teste da Boneca foi analisado qualitativamente com base na escolha de bonecas negras ou brancas em resposta a adjetivos positivos e negativos. P1, P2, P6 e P9 exibiram consistentemente viés antinegro no Teste da Boneca, escolhendo a boneca negra para todos os sete adjetivos negativos e a boneca

branca para todos os sete adjetivos positivos nas três linhas de base pré-intervenção. Esses participantes citaram a cor como justificativa para suas escolhas em pelo menos um dos ensaios de linha de base antes da intervenção. Por exemplo, P1 afirmou: “Eu não gosto dessa cor”; P2 mencionou que “Pessoas dessa cor geralmente são o vilão”; P6 expressou que a cor estava associada a coisas negativas; e P9 declarou que “Preto representa o mal e branco representa o bem”. É notável que todos os participantes demonstraram consistentemente um padrão binário de escolha, no qual selecionavam uma boneca para os adjetivos positivos e a outra para os adjetivos negativos.

Após a intervenção, P2, P6 e P9 não apresentaram mudanças e continuaram a selecionar a boneca negra para todos os adjetivos negativos e a boneca branca para todos os adjetivos positivos. No entanto, P1 mudou seu padrão de seleção e passou a escolher a boneca negra para adjetivos positivos. Em LB4, a participante selecionou a boneca negra para cinco adjetivos positivos (boa, inteligente, bonita, confiável e forte) e dois adjetivos negativos (irritante e inferior). Em LB5, a participante selecionou a boneca negra para quatro adjetivos positivos (boa, inteligente, confiável e forte) e três adjetivos negativos (irritante, feia e inferior). No follow-up, a boneca negra foi selecionada para cinco adjetivos positivos (boa, legal, bonita, forte e superior) e dois adjetivos negativos (perigosa e não inteligente). Ressalta-se que, quando questionada sobre o motivo de suas escolhas, a participante forneceu justificativas sem mencionar atributos raciais (por exemplo, “Eu escolhi porque gostei”, “Eu escolhi porque ela é boa”).

P3, P4 e P5 selecionaram consistentemente a boneca negra para todos os sete atributos positivos e a boneca branca para todos os sete atributos negativos

antes da intervenção. No entanto, esse padrão de escolha foi inconsistente com os dados obtidos na escala Likert e no teste relacional, nos quais foi identificado um viés racial negativo em relação a faces negras. Vale destacar que as justificativas fornecidas pelos participantes podem sugerir uma tentativa de ocultar vieses durante o teste. Por exemplo, P4 respondeu: “Eu não sou preconceituoso”, e por isso escolheu a boneca negra, enquanto P5 afirmou: “Não podemos ser racistas”. Após a intervenção, esses participantes não alteraram seus padrões de escolha. Já P7 e P8 apresentaram mudanças em seus padrões de seleção antes da intervenção e, por esse motivo, seus dados não foram incluídos na análise individual.

Quanto ao padrão de seleção dos participantes, analisamos os adjetivos mais frequentemente associados às bonecas negras e brancas. Observou-se que o adjetivo “forte” foi consistentemente atribuído à boneca negra por cinco participantes (55,55%) nas três linhas de base pré-intervenção. Por outro lado, os atributos “legal”, “superior”, “bonita” e “confiável” foram mais frequentemente selecionados para a boneca branca por cinco ou seis participantes nas três primeiras avaliações de linha de base. Após a intervenção, cinco participantes escolheram consistentemente a boneca branca para os adjetivos “legal”, “superior” e “inteligente”, enquanto cinco ou seis participantes selecionaram a boneca negra para os atributos “boa”, “forte” e “confiável”.

### **Discussão**

Este estudo teve como objetivo investigar um meio de alterar as funções avaliativas de estímulos socialmente relevantes por meio de um paradigma de

relações conflitantes. O emprego de múltiplas linhas de base antes e após a intervenção permitiu uma análise mais refinada de quaisquer mudanças observáveis no comportamento, tanto intra quanto entre participantes. Replicamos com sucesso os achados anteriores de Mizael et al. (2016, 2021), reduzindo a zero a frequência de faces negras relacionadas a estímulos negativos em um teste de relações por pareamento ao modelo (MTS) e atenuando as diferenças na valência das avaliações de faces brancas e negras. Além disso, 88,88% dos participantes formaram com sucesso classes de equivalência, apesar do uso de um paradigma de relações conflitantes que, teoricamente, deveria dificultar essa aquisição (ver de Carvalho & de Rose, 2014; Haydu et al., 2015; Watt et al., 1991). Esse resultado provavelmente se deve ao procedimento estendido e aos testes extensivos de simetria e transitividade durante a fase de treino (ver Adams et al., 1993; Arntzen & Holth, 1997; Mizael et al., 2016, 2020).

No teste relacional, reduzimos com sucesso o índice de viés racial para todos os participantes que formaram classes de equivalência. No entanto, apesar de eliminar a relação entre faces negras e o símbolo negativo, seis participantes continuaram a relacionar símbolos positivos com maior frequência às faces brancas (ver também Mizael et al., 2016, 2020). Esse padrão revela uma assimetria importante: a intervenção foi eficaz em reduzir relações negativas envolvendo faces negras, mas não em equiparar a frequência com que faces negras e brancas foram relacionadas a estímulos positivos, de modo que a desigualdade relacional persistiu em favor das faces brancas. Embora a maioria dos participantes tenha persistido em selecionar faces brancas quando acompanhadas de símbolos positivos, é importante destacar que, de acordo com

a definição de viés racial adotada no estudo, centrada em avaliações negativas (Mizael & de Rose, 2017; Nogueira, 2007), a diminuição da frequência de símbolos negativos relacionados às faces negras representa um resultado promissor. Ainda assim, estudos futuros poderiam explorar mudanças no procedimento que maximizem a probabilidade de relacionar faces negras a estímulos de valência positiva, e não apenas reduzir relações com estímulos negativos (e.g., Bierley et al., 1985; Bortoloti & de Rose, 2009). Uma mudança metodológica potencialmente promissora seria estabelecer relações derivadas entre faces brancas e estímulos neutros, e entre faces negras e estímulos positivos, a fim de modificar ainda mais as diferenças na avaliação de estímulos raciais (e.g., Santos & de Rose, 2019).

Na escala Likert, faces negras receberam predominantemente avaliações negativas no período pré-intervenção, enquanto faces brancas foram avaliadas entre ligeiramente negativas e neutras. No pós-intervenção, houve uma redução notável nas diferenças de valência entre faces negras e brancas, resultando em avaliações mais semelhantes, em consonância com achados anteriores (Mizael et al., 2016, 2020). É importante destacar, contudo, que essa aproximação não refletiu uma mudança nas avaliações das faces negras, que permaneceram essencialmente estáveis, mas sim uma redução nas avaliações positivas atribuídas às faces brancas. Em outros termos, a intervenção afetou colateralmente a valência das faces brancas sem alterar diretamente a valência das faces negras, o que representa um padrão distinto do esperado. Esse resultado foi inesperado, pois se esperava que as funções avaliativas dos estímulos positivos transformassem a valência das faces negras. No entanto, as

funções avaliativas negativas dos estímulos negativos pareceram transformar a valência das faces brancas. Esse achado ressalta a importância de pesquisas mais focadas em processos, uma vez que não é simples controlar precisamente os meios pelos quais contingências paralelas programadas produzem efeitos complementares. Por exemplo, pesquisas anteriores indicam que aumentos na medida geral de relacionamento entre itens de classes diferentes dependem mais da diminuição da fluência em relacionar estímulos não relacionados do que do aumento na relação entre estímulos relacionados (Cummins et al., 2018). Trata-se de dinâmicas complexas que só podem ser esclarecidas por meio de pesquisas sistemáticas contínuas, controlando uma ampla variedade de variáveis.

Este estudo representa a primeira pesquisa conhecida, no contexto de equivalência de estímulos e viés racial, que utiliza uma medida implícita tanto no pré-teste quanto no pós-intervenção. Na medida implícita, a maioria dos participantes apresentou efeitos FAST pequenos e não significativos, indicando ausência de viés. Ainda assim, em nível grupal, o viés ocorreu na direção esperada (antinegro/pró-branco) nas três linhas de base pré-intervenção. Após a intervenção, não foram encontradas evidências de alteração no viés antinegro na medida implícita: apesar de um leve aumento na fluência no bloco inconsistente, o padrão relacional identificado pelo FAST permaneceu essencialmente estável, sugerindo que a intervenção baseada em equivalência de estímulos não produziu efeitos detectáveis sobre o responder relacional implícito.

Esse resultado pode indicar a ineficácia do paradigma de relações conflitantes (CRP) em aumentar a relação entre faces negras e imagens positivas e/ou a insensibilidade do procedimento FAST para detectar viés racial. Antes de

tirar conclusões definitivas, contudo, é importante considerar que o FAST foi utilizado como medida de generalização, e a presença de viés implícito não era um critério de entrada no estudo. Para compreender melhor os efeitos do CRP sobre medidas implícitas, seria interessante selecionar apenas participantes com alto viés implícito antes da intervenção. Além disso, o número de participantes neste estudo foi menor do que em estudos FAST previamente publicados, e a variabilidade intra-indivíduos entre blocos podem ter sido grandes demais para que um efeito estatisticamente significativo fosse detectado com uma amostra tão pequena.

Ainda assim, a tendência consistente de efeitos em nível grupal na direção esperada (antinegro/pró-branco) ao longo de pelo menos cinco iterações é altamente sugestiva da presença de um viés pequeno, ainda que estatisticamente não significativo. Essa tendência de desempenho diferencial entre blocos do FAST não foi alterada pelo CRP, o que pode sugerir que a eficácia dessas intervenções para modificar escores de viés implícito não foi demonstrada. Embora conclusões definitivas sejam frágeis na ausência de significância estatística robusta, essa interpretação é compatível com pesquisas que indicam que vieses explícitos tendem a mudar antes dos implícitos, uma vez que estes últimos se baseiam em padrões históricos de resposta com grande momentum comportamental, enquanto respostas em medidas explícitas são mais fortemente controladas por contingências locais (ver Hughes & Barnes-Holmes, 2011; Lai et al., 2014, 2016). Estudos futuros deveriam investigar outros procedimentos para reduzir o viés racial implícito, considerando as implicações dessas reduções em um contexto social mais amplo, como a seleção de pares em ambientes sociais.

Dos nove participantes que inicialmente apresentaram viés racial no teste relacional, quatro também demonstraram viés no Teste da Boneca. Após a intervenção, apenas um participante apresentou mudança no viés nesse teste, sugerindo uma possível generalização dos efeitos da intervenção para medidas mais próximas do ambiente natural da criança. Contudo, essa mudança ocorreu em apenas um participante, indicando evidência limitada de generalização. A dificuldade em generalizar a transformação de funções de estímulos para o Teste da Boneca, que possui uma topografia distinta dos estímulos da intervenção, pode exigir maior variabilidade de estímulos durante os procedimentos. Além disso, três crianças apresentaram um viés favorável às bonecas negras no Teste da Boneca, em contraste com seu desempenho no teste relacional, o que levanta questionamentos sobre a confiabilidade do Teste da Boneca e a possibilidade de respostas socialmente desejáveis (Carifio, 1994; Klesges et al., 2004).

Apesar das preocupações quanto à validade e à evidência limitada de generalização no Teste da Boneca, vale destacar alguns padrões de resposta interessantes. Antes da intervenção, o atributo “forte” foi consistentemente associado às bonecas negras por cinco dos nove participantes (55,55%) nas três avaliações de linha de base. Em contraste, atributos como “legal”, “bonita”, “confiável” e “superior” foram mais frequentemente atribuídos às bonecas brancas. O adjetivo “forte”, embora favorável em muitos contextos, pode carregar conotações negativas e estereótipos raciais (Conrado & Ribeiro, 2017), como a ideia de que pessoas negras são mais adequadas ao trabalho manual. Assim, a inclusão desse adjetivo pode ter criado ambiguidade nas opções de resposta. Após a intervenção, houve uma mudança nas escolhas: cinco participantes

passaram a selecionar a boneca negra para atributos como “boa” e “confiável”, e “forte” deixou de ser o adjetivo mais associado às bonecas negras. Esses resultados sugerem que a intervenção alterou os padrões avaliativos dos participantes e sustentam a interpretação de que o adjetivo “forte” pode ter sido utilizado de forma negativa em pelo menos alguns casos.

Uma limitação do estudo foi a baixa correspondência entre as medidas usadas para seleção dos participantes (teste relacional) e as medidas utilizadas para investigar a generalização do procedimento (Teste da Boneca, FAST). Na literatura da cognição social, não se espera que vieses explícitos e implícitos se correlacionem fortemente (e.g., Dasgupta & Greenwald, 2001; Greenwald et al., 1998; Payne, 2001; Sawyer & Gampa, 2018), mesmo considerando apenas diferenças metodológicas e de estímulos (Payne et al., 2008). Como os testes implícitos foram originalmente concebidos para medir vieses automáticos fora da consciência (Banaji & Greenwald, 2013), a ausência de correlação com medidas explícitas não é surpreendente. Ainda assim, a ausência de efeitos pré-intervenção significativos no FAST pode indicar insensibilidade do instrumento nessa população, possivelmente devido à dificuldade da tarefa ou à inadequação dos estímulos para crianças. São necessários mais estudos para validar a sensibilidade do FAST ao viés racial infantil, enfatizando a interpretação cuidadosa de escores individuais e efeitos em nível grupal, especialmente em amostras pequenas (Cummins & Roche, 2020).

A falta de correspondência entre o teste relacional e o Teste da Boneca, ambos supostamente medindo viés racial, questiona a confiabilidade do Teste da Boneca e aspectos procedimentais. Ambos os testes avaliam estímulos diferentes

e podem não medir diretamente o viés racial, mas inferi-lo indiretamente. Esperar correlação entre eles pressupõe a existência de um único construto subjacente, o que pode ser excessivamente ambicioso. A generalização e os efeitos globais da intervenção sobre esse construto permanecem especulativos. Intervenções mais abrangentes e de maior duração podem ser necessárias para produzir efeitos generalizados em diferentes medidas de viés racial. Além disso, uma especificação procedimental mais rigorosa do Teste da Boneca poderia reduzir a variabilidade entre medidas. Por exemplo, a posição da pesquisador durante o teste pode ter influenciado as respostas ao tornar pistas sutis mais salientes, aumentando características de demanda (Krumpal, 2013; Rosnow & Rosenthal, 1976). Ademais, o formato das perguntas pode ter limitado as opções de resposta, contribuindo para a baixa variabilidade ao longo das avaliações.

Outra limitação foi a ausência de uma medida de validade social da intervenção.

Estudos futuros poderiam incluir avaliações com pais, professores e gestores escolares para investigar se os efeitos da intervenção impactam as relações cotidianas dos participantes. Além disso, seria valioso empregar novas medidas para examinar a generalização do procedimento em ambientes mais naturalísticos, como a observação das interações das crianças na escola. Cabe destacar ainda que a pesquisadora permaneceu presente na sala durante a aplicação da intervenção, o que pode ter funcionado como variável contextual adicional, alterando o padrão de respostas por meio de controle por audiência. Estudos futuros deveriam avaliar os efeitos da intervenção em condições de maior autonomia do participante, com coleta automatizada e sem a presença do experimentador, a fim de isolar os efeitos das contingências programadas.

## Direções Futuras

O preconceito racial é um fenômeno complexo e multifacetado (Matsuda et al., 2020; Parks & Kirby, 2022; Passarelli et al., 2023). Este estudo concentrou-se no aspecto avaliativo, empregando medidas para investigar a relação entre estímulos raciais e estímulos de valência positiva ou negativa. A intervenção foi bem-sucedida em mitigar avaliações negativas direcionadas a faces negras (ver também Mizael et al., 2016, 2020). No entanto, ainda é necessária maior clareza quanto à generalização desse procedimento para ambientes mais naturalísticos.

Pesquisas futuras deveriam investigar medidas para identificação do preconceito racial e para a avaliação da generalização das intervenções. Para além da valência dos estímulos raciais, outras propriedades dos estímulos também devem ser exploradas nessa linha de pesquisa, como relações hierárquicas entre estímulos e preferências por estímulos. Diferentes tarefas de tomada de decisão envolvendo estímulos também podem ser empregadas como medidas (e.g., Correll et al., 2006; *shooter game*). Além disso, a avaliação das relações entre estímulos em termos de relações diferentes da equivalência (e.g., “mais inteligente que” ou atratividade), utilizando procedimentos como o IRAP (ver De Houwer et al., 2015; Murphy et al., 2015), permitiria tanto a avaliação quanto a modificação de estereótipos mais sofisticados.

Uma metodologia promissora que pode servir de base para intervenções deriva de conceitos da Teoria dos Quadros Relacionais (Relational Frame Theory – RFT; Hayes et al., 2001) e é consistente com a Terapia de Aceitação e Compromisso (Acceptance and Commitment Therapy – ACT; Hayes et al., 2013). No arcabouço da ACT, o preconceito racial pode ser compreendido como uma

forma de fusão cognitiva, na qual relações verbais excessivamente rígidas associadas ao viés racial foram estabelecidas e operam com considerável momentum. O correlato da fusão cognitiva é a inflexibilidade psicológica (Florez et al., 2019; Flynn et al., 2018; Levin et al., 2016). A ACT desenvolveu uma série de procedimentos voltados a aumentar a flexibilidade psicológica e reduzir a fusão cognitiva, os quais podem ter utilidade mensurável em intervenções sobre viés racial (Hayes et al., 2004; Lillis & Hayes, 2007; Masuda et al., 2007). Por exemplo, Lillis e Hayes (2007) demonstraram que uma intervenção baseada em ACT, focada no reconhecimento de pensamentos preconceituosos e na clarificação de valores relacionados a outras raças, resultou em aumento de intenções comportamentais positivas em estudantes universitários.

De fato, alguns autores sugeriram que uma possível intervenção para o viés racial poderia empregar procedimentos de *defusão* cognitiva (Holmes et al., 2024; Matsuda et al., 2020). Uma estratégia potencial para promover a *defusão* cognitiva envolve a repetição de palavras ou frases, com o objetivo de reduzir as funções aversivas dos estímulos ao aumentar a discriminação de suas características topográficas (ver Kishita et al., 2014). Nesse sentido, um indivíduo com viés contra pessoas negras poderia ser treinado a formar relações nas quais estímulos raciais negros sejam relacionados a estímulos negativos em um contexto, mas também a estímulos positivos em outro contexto. Tal estratégia ampliaria o repertório de funções atribuídas ao estímulo problemático e aumentaria o leque de probabilidades de resposta para além de padrões como agressão ou esquiva.

É interessante notar que uma abordagem relativamente recente na literatura da cognição social, semelhante às técnicas de *defusão cognitiva*, foi

utilizada para alterar o viés racial implícito (Lai et al., 2014). Em particular, nesse estudo, Lai et al. (2014) possivelmente aplicaram inadvertidamente uma intervenção de defusão cognitiva ao expor repetidamente os participantes a um teste implícito de viés racial (Implicit Association Test – IAT; Greenwald et al., 1998). A proposta era aumentar a fluência de desempenho em ambos os blocos do teste por meio da exposição repetida e, assim, reduzir parcialmente a preferência que muitos participantes apresentam por responder de forma congruente a estímulos negros e negativos e a estímulos brancos e positivos, em vez do inverso. Sob uma perspectiva comportamental, esse procedimento funcionaria ao opor-se ao momentum comportamental de respostas relacionais previamente estabelecidas ao longo de um número suficiente de sessões de prática envolvendo contingências opostas. No estudo de Lai et al. (2014), o método de exposição repetida ao teste implícito mostrou-se um dos mais eficazes entre 17 intervenções para reduzir o viés racial, resultado posteriormente confirmado por Lai et al. (2016).

Uma observação importante do estudo de Lai et al. (2014) é que as oito intervenções laboratoriais mais eficazes para reduzir o viés racial compartilhavam algumas características comuns. Em geral, os métodos bem-sucedidos envolviam a apresentação direta de estímulos raciais juntamente com outros estímulos evocativos. Por exemplo, intervenções de condicionamento avaliativo, que pareiam imagens de faces negras e brancas com estímulos positivos e negativos evocativos, resultaram na redução de avaliações negativas explícitas e implícitas de faces negras (Lai et al., 2014). No entanto, um estudo subsequente (Lai et al., 2016) mostrou que o condicionamento avaliativo isoladamente pode não reduzir

de forma consistente o viés racial implícito, uma vez que sua eficácia foi demonstrada principalmente em amostras grandes e utilizando tarefas implícitas do tipo *go/no-go* como medida dependente.

O presente estudo pode ser considerado complementar a essa estratégia mais ampla de modificação de associações entre estímulos, embora aqui tal modificação tenha sido obtida por meio de um paradigma de relações derivadas, e não de condicionamento avaliativo. Como demonstrado por Lai et al. (2016), é possível que o uso isolado do paradigma de relações conflitantes (CRP) não seja suficiente para reduzir o preconceito racial em medidas de generalização, indicando a necessidade de combinar diferentes procedimentos. Além disso, os resultados de intervenções de condicionamento avaliativo sobre viés racial são difíceis de generalizar para estudos com amostras menores e com maior diversidade de medidas de desfecho. Esses achados ressaltam a complexidade do desenvolvimento de intervenções laboratoriais para a redução do preconceito racial. Ainda assim, espera-se que o estudo aqui apresentado ofereça aos pesquisadores abordagens metodológicas viáveis para explorar o desenvolvimento dessas intervenções, bem como evidencie a utilidade potencial de novas medidas de desfecho.

-----

**Comentário: Replicar não é apenas repetir efeitos, mas identificar limites e gerar novas questões de pesquisa**

A replicação foi bem-sucedida em um aspecto central: a formação de classes de equivalência envolvendo faces negras e estímulos positivos ocorreu para a maioria dos participantes e foi mantida após 15 dias. Como consequência direta dessa reorganização, a frequência de responder relacional entre faces negras e estímulos negativos foi eliminada no teste relacional, replicando com precisão o efeito descrito em estudos anteriores. Além disso, a intervenção foi bem-sucedida em mitigar diferenças na valência entre faces brancas e negras. Em outras palavras, antes do procedimento, faces brancas eram avaliadas como significativamente mais positivas do que faces negras; após a intervenção, essa diferença deixou de ser observada. Esse resultado reafirma que o paradigma de relações conflitantes pode, sob parâmetros adequados, superar histórias pré-experimentais socialmente estabelecidas, ao menos em contextos controlados.

No entanto, a própria topografia da mudança observada levanta novas questões. A redução do viés não se deu principalmente pelo aumento substancial da relação entre faces negras e estímulos positivos, mas pela diminuição da relação entre faces negras e estímulos negativos. Em termos funcionais, o procedimento mostrou-se mais eficaz em enfraquecer funções aversivas do que em estabelecer funções positivas. Esse padrão sugere que reduzir o viés pode não ser funcionalmente equivalente a promover avaliações positivas e que esses dois efeitos podem depender de variáveis distintas.

As medidas de generalização ampliaram ainda mais esse quadro e introduziram um desafio interpretativo importante. As medidas utilizadas para selecionar os participantes para a intervenção não apresentaram alta correspondência funcional com aquelas empregadas para avaliar generalização, mesmo antes da intervenção. Em outras palavras, as

crianças foram selecionadas com base em padrões estáveis de viés identificados em medidas relacionais e avaliativas específicas, enquanto a generalização foi examinada por meio de medidas que envolviam diferentes topografias de resposta e diferentes controles contextuais. Essa baixa correspondência entre critérios de seleção e medidas de generalização impôs limites claros à interpretação direta dos efeitos interventivos.

Nesse ponto, emerge uma questão central: o problema empírico não é se o viés foi reduzido, mas em quais contextos essa redução se manifesta e quais formas de responder enviesado permanecem inalteradas. Os resultados indicam que o viés racial não se apresenta de forma uniforme entre diferentes medidas, nem responde de maneira homogênea às tentativas de intervenção.

Dentro desse cenário, mudanças consistentes foram observadas no teste relacional e na escala Likert, ambas funcionalmente mais próximas do objetivo do procedimento interventivo, isto é, da reorganização de avaliações explícitas e relações simbólicas diretamente treinadas ou testadas em condições de escolha controlada. Nessas medidas, os efeitos foram claros e replicáveis para a maioria dos participantes selecionados. Em contraste, os efeitos foram substancialmente mais limitados nas medidas menos isomórficas ao procedimento de intervenção.

No FAST, apenas três participantes apresentavam viés implícito antinegro claro antes da intervenção, e as mudanças observadas após o treino foram instáveis. No Teste da Boneca a generalização ocorreu em apenas um dos quatro participantes que inicialmente apresentavam viés racial antinegro nessa medida. Os cinco participantes remanescentes exibiram, desde o pré-teste, respostas fortemente moduladas por deseabilidade social, frequentemente acompanhadas de verbalizações como “a gente não pode ser

preconceituoso” ou “eu não sou racista”, independentemente do padrão relacional identificado em outras tarefas.

Essa baixa correspondência entre medidas de seleção e medidas de generalização impõe novas questões: se estamos medindo o mesmo fenômeno, por que as medidas não correspondem? Uma das hipóteses levantadas para o Teste da Boneca Adaptado diz respeito ao viés de desejabilidade social. Um aspecto do procedimento merece destaque: durante a aplicação do teste, o experimentador permanecia de frente para a criança, o que pode ter contribuído para a emissão de respostas socialmente desejáveis. Nesse contexto, emerge uma questão adicional: crianças mais jovens apresentariam maior convergência entre as medidas? Como discutido no capítulo anterior, crianças nessa faixa etária possivelmente operam sob menor controle de normas sociais explícitas, o que poderia reduzir esse tipo de filtro social.

Um dado que merece atenção particular diz respeito aos participantes que não formaram classes de equivalência ao longo do procedimento. Em uma criança, apesar da exposição completa ao treino, não se observou a emergência consistente das relações derivadas esperadas. De forma sistemática, esse mesmo participante também não apresentou mudanças nas medidas de viés racial utilizadas no pré e pós-teste. Essa convergência negativa não é trivial. Ela sugere que a ausência de mudança não pode ser atribuída simplesmente à insensibilidade das medidas ou à variabilidade aleatória do responder, mas está diretamente relacionada à falha na reorganização das relações condicionais treinadas. Trata-se, evidentemente, de uma observação preliminar, cuja relevância dependerá de sua replicação em estudos subsequentes, mas que já se apresenta como um ponto empírico digno de atenção.

Dessa forma, novas questões permanecem em aberto. Já não se trata apenas de saber se é possível reduzir o preconceito racial por meio de procedimentos baseados em equivalência de estímulos. Os estudos desta tese demonstraram efeitos consistentes no teste relacional e na escala de valência, mas com generalização parcial e variável para outras medidas, o que impede afirmações mais amplas sobre replicabilidade. Essa divergência entre medidas levanta uma questão metodológica adicional: em que medida os efeitos observados refletem reorganização relacional genuína e em que medida são modulados pela presença da pesquisadora como variável contextual? A audiência do experimentador funciona como estímulo discriminativo que sinaliza contingências sociais de aprovação e punição, podendo selecionar respostas socialmente desejáveis independentemente das contingências experimentais programadas. Evidências nessa direção foram documentadas por Rogers e Prentice-Dunn (1981), que demonstraram que condições de anonimato produziram padrões de resposta racialmente enviesados substancialmente distintos daqueles observados em condições de identificação. Replicações futuras deveriam examinar sistematicamente os efeitos do anonimato, por meio de coleta automatizada, sem presença do experimentador e com garantia explícita de que as respostas individuais não serão identificáveis, para separar o controle exercido pelas contingências relacionais programadas do controle exercido pela audiência social durante a aplicação.

A questão desloca-se, portanto, para quando, para quem e sob quais condições esses efeitos se mostram mais amplos e funcionalmente relevantes. Em particular, torna-se central investigar se, em crianças mais jovens, recém-ingressas no contexto escolar e com uma história pré-experimental potencialmente menos densa e menos rigidamente verbal, haveria maior correspondência entre diferentes medidas de viés racial. Essa condição poderia favorecer não apenas a formação de classes de equivalência, mas também uma

generalização mais consistente entre avaliações relacionais, medidas implícitas e respostas em contextos socialmente mais naturais. Além disso, emerge a necessidade de examinar a validade social dessas mudanças, isto é, como alterações observadas em tarefas experimentais se traduzem em padrões de resposta com maior relevância para interações cotidianas. É esse conjunto de questões que orienta o estudo apresentado a seguir.

\*\*\*\*\*

**Estudo 5 - Efeitos da Formação de Classes de Equivalência na Redução do  
Preconceito Racial em Crianças I: Evidências de Generalização e  
Manutenção <sup>10</sup>**

Pesquisas anteriores demonstraram que o estabelecimento de relações de equivalência entre faces negras e símbolos positivos reduz avaliações negativas dessas faces. Este estudo replicou esses achados e investigou sua generalização e manutenção. Participaram 18 crianças de 6 anos (11 meninos), que inicialmente relacionaram faces negras a símbolos negativos em um teste relacional de pareamento ao modelo (MTS) e as avaliaram negativamente em uma escala Likert. As crianças foram treinadas a relacionar um símbolo abstrato (B1) a um símbolo positivo (A1) e, posteriormente, B1 a faces negras (C1). Utilizou-se um delineamento de sondagens múltiplas não concorrentes, com follow-up realizado quinze dias após a intervenção. Todos os participantes formaram classes de equivalência entre faces negras e o símbolo positivo e deixaram de relacionar essas faces ao símbolo negativo no teste MTS. Após a formação das classes, as avaliações de faces negras tornaram-se mais positivas na escala Likert, sendo mantidas pela maioria dos participantes no follow-up. Além disso, essas mudanças generalizaram para o teste das bonecas, no qual as escolhas da boneca negra diminuíram diante de adjetivos negativos e aumentaram diante de adjetivos positivos, embora alguns participantes não tenham apresentado ou mantido esse efeito. Os resultados replicam achados anteriores com um delineamento mais robusto, demonstrando a transferência de funções positivas para diferentes medidas avaliativas.

---

<sup>10</sup> Passarelli, D. A., Mizael, T.M.M., Roche, B., De Rose, J.C. (2026). Efeitos da Formação de Classes de Equivalência na Redução do Preconceito Racial em Crianças I: Evidências de Generalização e Manutenção. [Manuscrito premiado na *Student Paper Competition*, do *Special Interest Group* em *Experimental Analysis of Human Behavior (EAHB sig)*; a ser submetido a periódico científico].

## Introdução

O preconceito racial pode ser definido como atitudes negativas em relação a um indivíduo ou a um grupo específico. Essas atitudes são influenciadas pela cultura e baseiam-se em características fenotípicas, como cor da pele, textura do cabelo ou traços faciais (Nogueira, 2006; Mizael & de Rose, 2017; Passarelli et al., 2023). Especificamente, o viés antinegro refere-se a atitudes negativas direcionadas a pessoas negras ou a símbolos que representam uma origem afrodescendente. A literatura apresenta diferentes interpretações acerca da conceitualização de atitudes (ver Mizael & de Rose, 2017; Bohner & Dickel, 2011; Gawronski & LeBel, 2008; Guerin, 2005). Na psicologia experimental, definir atitudes como avaliações positivas ou negativas de um estímulo relevante facilita a mensuração desse fenômeno.

O viés antinegro pode manifestar-se por meio de avaliações explícitas, caracterizadas por respostas relacionais extensas e complexas, emitidas sem restrições temporais para sua expressão (Barnes-Holmes et al., 2010). Um exemplo é afirmar abertamente: “Eu não gosto de pessoas negras porque elas geralmente são preguiçosas”. Diversos métodos podem ser utilizados para avaliar essas atitudes explícitas, incluindo questionários, escalas Likert (Bradley & Lang, 1994) ou a exposição a estímulos que representam uma raça — como bonecas negras e faces negras — no contexto de estímulos positivos ou negativos (Clark & Clark, 1947; Mizael et al., 2016, 2020). Uma forma de alterar atitudes preconceituosas consiste em modificar a avaliação de estímulos raciais. Alguns estudos investigaram a modificação de atitudes por meio do pareamento de estímulos-alvo, como imagens de faces, com símbolos positivos, com o objetivo

de transferir propriedades positivas a esses estímulos e, conseqüentemente, alterar as avaliações negativas a eles “associadas” (Hofmann et al., 2010; Lai et al., 2016; Olson & Fazio, 2001).

No paradigma de equivalência de estímulos, é possível estabelecer relações indiretas entre estímulos por meio de tarefas de discriminação condicional. Esse paradigma permite que novas relações emergjam a partir do treino direto de algumas relações. Quando as relações emergentes demonstram reflexividade, simetria e transitividade das relações treinadas, os estímulos relacionados são considerados como formando uma classe de estímulos equivalentes, e as funções de um estímulo podem transferir-se para os demais membros da classe (e.g., de Rose et al., 1988; Dougher et al., 1994; Martins et al., 2023). Por exemplo, ao treinar relações de pareamento ao modelo (MTS) entre um estímulo A (símbolo positivo) e um estímulo B (estímulo abstrato), e entre B e um estímulo C (faces negras), é possível que A, B e C formem uma classe de estímulos equivalentes, e que as funções do símbolo positivo sejam transferidas para as faces negras (Mizael et al., 2016, 2020; Passarelli, Roche, et al., 2024).

A literatura indica que o paradigma de equivalência de estímulos tem sido uma ferramenta eficaz para a modificação de avaliações (e.g., Barnes-Holmes et al., 2000; Hughes et al., 2018; Mizael et al., 2016, 2020; Ribeiro et al., 2020; Santos & de Rose, 2018, 2019; Smeets & Barnes-Holmes, 2003). Em um estudo conduzido por Santos e de Rose (2018), crianças em idade pré-escolar foram treinadas a relacionar estímulos abstratos a personagens favoritos (A1) ou não apreciados (A2), formando classes de equivalência com formas geométricas (B1 e B2) e estímulos abstratos (C1 e C2). Posteriormente, as crianças receberam dois

pedaços do mesmo alimento, rotulados, respectivamente, com C1 (equivalente ao personagem favorito) e C2 (equivalente ao personagem não apreciado). A maioria das crianças considerou que o alimento rotulado com C1 era mais saboroso do que aquele rotulado com C2.

Um estudo conduzido por Mizael et al. (2016) treinou treze crianças com atitudes negativas em relação a faces negras a relacionar um estímulo abstrato (B1) a um estímulo positivo (A1) e, em seguida, o estímulo abstrato (B1) a faces negras. Após o treino, as crianças formaram classes de equivalência relacionando faces negras a estímulos positivos, o que resultou em uma redução estatisticamente significativa na diferença entre as avaliações de faces brancas e negras, bem como em uma diminuição da frequência com que faces negras eram associadas a um símbolo negativo em um teste relacional de pareamento ao modelo. Esses resultados foram replicados por Mizael et al. (2020) e, mais recentemente, por Passarelli et al. (2024).

O estudo de Passarelli et al. (2024) replicou o procedimento de Mizael e colaboradores (2016), mas incluiu medidas repetidas, permitindo verificar se a mera exposição aos procedimentos de mensuração produziria mudanças não relacionadas à formação de equivalência. Além disso, o estudo utilizou um conjunto mais variado de medidas para avaliar a generalidade dos efeitos. Nove crianças (M = 8,2 anos; cinco meninos) que atribuíram com maior frequência um símbolo negativo a faces negras e/ou um símbolo positivo a faces brancas participaram de três avaliações pré-teste consecutivas, compostas por quatro medidas distintas. Em seguida, os participantes foram treinados a estabelecer classes de equivalência entre faces negras e um estímulo positivo. Após o

procedimento, duas avaliações pós-teste (idênticas às pré-testes) com as quatro medidas avaliativas foram realizadas, seguidas por um teste de manutenção quinze dias após a formação das classes.

Oito participantes formaram classes de equivalência e mantiveram essas classes quinze dias após a intervenção. Após a formação das classes, esses participantes reduziram a frequência com que relacionavam faces negras a símbolos negativos para valores próximos de zero no teste relacional, e a diferença entre as avaliações de faces brancas e negras na escala Likert deixou de ser significativa. Um participante não formou classes de equivalência e continuou a relacionar faces negras a símbolos negativos e a avaliar faces negras de forma mais negativa do que faces brancas na escala Likert. Esses achados replicam os resultados obtidos por Mizael et al. (2016, 2020) e demonstram que a mera exposição aos procedimentos de mensuração não é suficiente para reduzir atitudes negativas em relação a faces negras. Os resultados também indicam que os efeitos da intervenção nessas duas medidas (escala Likert e teste relacional) foram mantidos em uma avaliação de seguimento quinze dias após a intervenção.

Entretanto, os resultados não foram consistentes para duas medidas adicionais coletadas no estudo. O FAST (Functional Acquisition Speed Test; O'Reilly et al., 2012), uma medida implícita, e o Teste da Boneca (Clark & Clark, 1947), no qual os participantes relacionavam uma boneca negra e uma branca a atributos positivos e negativos, foram utilizados para explorar a relação entre as relações treinadas e a generalização do treino discriminativo para estímulos novos. Os resultados apresentaram menor estabilidade nas linhas de base pré-intervenção, indicando variabilidade entre os participantes. Especificamente no

Teste da Boneca, o viés racial antinegro foi observado em quatro participantes, enquanto a generalização ocorreu apenas para um participante. Evidências parciais de generalização foram observadas somente na escala Likert e no teste relacional. Cabe destacar, contudo, que a redução das diferenças na escala Likert não decorreu de um aumento nas avaliações positivas de faces negras, mas de uma diminuição nas avaliações positivas de faces brancas. Essa distinção tem implicações sociais relevantes: do ponto de vista da equidade racial, reduzir a avaliação positiva de faces brancas é um resultado qualitativamente distinto de aumentar a avaliação positiva de faces negras, e essa diferença não deve ser obscurecida pela apresentação exclusiva de dados quantitativos agregados.

O presente estudo buscou abordar limitações identificadas no trabalho de Passarelli, Roche et al. (2024). Devido à baixa correspondência entre as medidas de identificação do viés antinegro, foram realizadas adaptações na escala Likert e no Teste das Bonecas (ver Procedimento). Além disso, algumas respostas no teste das bonecas sugeriram que as crianças poderiam ter ocultado seus vieses raciais para fornecer respostas socialmente aceitáveis (Krumpal, 2013). Esperava-se que crianças não apresentassem sofisticação suficiente para ocultar vieses (Levy & Hughes, 2009). Por essa razão, a idade das crianças foi considerada uma variável importante a ser controlada no novo estudo, que foi conduzido com crianças de 6 anos de idade. Foi implementado um Delineamento de Sondas Múltiplas Não Concorrente entre participantes (Cooper et al., 2020). As medidas utilizadas foram as mesmas empregadas por Passarelli et al. (2024), com exceção da medida implícita, considerada complexa demais para crianças de 6 anos. As crianças participaram de avaliações pré e pós-intervenção, com o número de

sondas variando entre participantes. Adicionalmente, professores responderam a um questionário trinta dias após a intervenção, contendo questões relacionadas à validade social do procedimento empregado.

## Método

### Participantes

Setenta e uma crianças (M = 6 anos; 36 do sexo masculino) de uma escola estadual de São Carlos foram avaliadas por meio de um Teste Relacional e de uma Escala *Likert* (ver Local e Instrumentos, abaixo). Dezoito crianças (M = 6 anos; 11 do sexo masculino) que apresentaram avaliações negativas em relação a faces negras em ambos os testes foram selecionadas para participar do estudo. A classificação étnico-racial dos participantes foi realizada pelos pais ou responsáveis durante o preenchimento do Termo de Consentimento, sendo 11 participantes (61,11%) classificados como brancos e 7 (38,89%) como pardos. A Tabela 1 apresenta características dos participantes

### Tabela 1

#### *Caracterização dos participantes*

Identificação	Gênero	Idade	Raça	Ano	N (%)
P1, P2, P4, P6, P8, P9, P12, P15 P16	Masculino	6 anos	Branco	1°	9(50%)
P3, P18	Feminino	6 anos	Branco	1°	2(11.1%)
P14, P5	Masculino	6 anos	Pardo	1°	2(11.1%)
P7, P10, P11, P13, P17	Feminino	6 anos	Pardo	1°	5(27.7%)

## Local e Estímulos

Os dados foram coletados na sala de leitura de uma Escola Estadual localizada no interior do Estado de São Paulo. Foi utilizado um computador Dell G15 Intel Core i5, 2.50 GHz, 8Gb de Memória, com os softwares MTS - III (Wallace, 2003) e Psychopy3.0 (Pierce et al., 2019) para apresentação dos estímulos e das consequências. Uma parte dos estímulos utilizados foram os mesmos dos estudos de Mizael e colaboradores (2016, 2020) e de Passarelli, Roche et al. (2024): A1 era uma mão fazendo sinal positivo, A2 era uma mão fazendo sinal negativo. B1, B2 e C2 eram estímulos abstratos. C1 consistia em quatro faces negras (duas masculinas e duas femininas), enquanto C3 consistia em faces brancas (duas femininas e duas masculinas). Adicionalmente, foram utilizadas quatro faces com tonalidades de pele intermediárias (na escala Likert) e duas bonecas com características físicas idênticas, diferindo exclusivamente na tonalidade da pele (uma com pele clara e outra com pele escura), no Teste da Boneca.

## Instrumentos

*Teste Relacional.* Desenvolvido por Mizael et al. (2016), este teste permite analisar o padrão de responder relacional por meio do procedimento de MTS. Os modelos utilizados foram os símbolos positivo (A1) e negativo (A2), enquanto os estímulos de comparação consistiam em faces negras (C1), um estímulo abstrato (C2) e faces brancas (C3). O teste consistiu em 16 tentativas, sem consequências diferenciais, onde o participante foi instruído a clicar no estímulo de comparação que ele considerava correspondente ao modelo. Foi usada a seguinte fórmula para o cálculo do viés racial:

$$v = [(B+) - (N+)] + [(N-) - (B-)],$$

onde  $v$  representa o índice de viés.  $B+$  e  $N+$  são, respectivamente, o número de escolhas das faces brancas e negras diante do símbolo positivo, e  $B-$  e  $N-$  são, respectivamente, o número de escolhas das faces brancas e negras diante do símbolo negativo. O valor de  $v$ , portanto, aumenta quando o símbolo positivo é relacionado mais vezes com a face branca e quando o símbolo negativo é relacionado mais vezes com a face negra, podendo variar entre  $-16$  (faces brancas sempre relacionadas com o símbolo negativo e negras sempre com o positivo) a  $+16$  (faces negras sempre relacionadas com símbolo negativo e as faces brancas com o positivo)

*Escala Likert.* A escala foi programada no software *PsychoPy 3.0*. Os participantes avaliaram quatro faces negras e quatro faces brancas, sendo metade de cada grupo composta por faces femininas e metade por faces masculinas, apresentadas uma a uma. Uma escala deslizante de cinco pontos era exibida na tela, com um emoji abaixo de cada número representando emoções positivas (feliz), neutras ou negativas (triste). O objetivo do teste foi avaliar a valência atribuída às faces negras e brancas. Todas as imagens faciais foram obtidas do banco *FaceResearch* ([faceresearch.org](http://faceresearch.org)), que disponibiliza imagens para fins de pesquisa. Diferentemente de Passarelli et al. (2024), apenas faces foram apresentadas, sem o uso de desenhos animados como estímulos distratores, uma vez que havia evidências de que esses distratores enviesavam a avaliação de ambas as faces em direção à valência negativa.

*Teste da Boneca Adaptado (Clark & Clark, 1947).* Foram utilizadas duas bonecas, uma branca e uma negra. Foi elaborado um roteiro semiestruturado com doze questões avaliativas e duas questões de identificação. As questões avaliativas

contaram com seis adjetivos positivos e seis adjetivos negativos, de modo que as questões foram: Alguma é boa/má? Alguma é chata/legal? Alguma é inteligente/burra? Alguma é bonita/feia? Alguma é confiável/perigosa? Alguma é forte/fraca? As questões de identificação foram: Alguma é Branca/Negra? Alguma se parece com você? Para analisar o padrão de resposta do participante, foi realizado o mesmo cálculo do índice  $v$ , que foi utilizado no Teste Relacional. Um índice  $v$  positivo, indica que o adjetivo positivo controla mais frequentemente a escolha da boneca branca e o adjetivo negativo a escolha da boneca negra.

### **Procedimento**

O procedimento foi aprovado pelo Comitê de Ética em Pesquisa da Universidade Federal de São Carlos (CAAE nº 48028721.0.0000.5504) e pela Secretária Estadual de Educação do Estado de São Paulo. A participação foi voluntária, com consentimento dos responsáveis legais e assentimento das crianças, sendo os estudos conduzidos mediante autorização da Secretaria Pública de Educação e da direção das escolas envolvidas. Foi utilizado um Delineamento de Sondas Múltiplas Não Concorrente entre participantes (Cooper et al. 2020). Em cada sonda, foram coletadas três medidas avaliativas, na seguinte ordem:

*Teste relacional* conduzido por meio de um procedimento de *matching-to-sample* (MTS), envolvendo os conjuntos A (símbolos) e C (faces e figura abstrata). O modelo era apresentado no topo-central da tela, podendo ser um símbolo positivo (A1) ou negativo (A2). Inicialmente, o experimentador se certificou que o participante compreendia o significado dos símbolos A1 e A2. Em seguida, o participante era instruído a clicar no modelo, de modo que após clicar, apareciam três comparações:

uma face negra (C1), um estímulo abstrato (C2) e uma face branca (C3). A ordem de apresentação dos modelos e das comparações foi randomizada ao longo das 16 tentativas. Após as comparações aparecerem na tela, o participante era instruído a selecionar a comparação que ele considerasse correspondente ao modelo. O índice de viés racial varia de -16 (viés positivo em relação às faces negras) e 16 (viés negativo em relação às faces negras). Foram selecionados participantes com índice de viés maior que 6 em todas as sondagens antes da intervenção.

*Escala Likert*, onde a face era apresentada no topo central da tela e abaixo era apresentada uma escala de cinco pontos, com emojis que representavam emoções negativas, neutra ou positivas. Os participantes eram instruídos a atribuir uma avaliação às faces, de modo que 1-2 pontos seriam considerados avaliações negativas, 3 pontos, avaliação neutra, e 4-5 pontos, avaliações positivas. Os estímulos foram apresentados um de cada vez e a ordem de apresentação foi randomizada. O critério para presença de viés racial foi atribuir para as faces negras uma avaliação negativa (entre 1 e 2) e inferior à atribuída as faces brancas.

*Teste da Boneca*. O participante se sentou em uma mesa, onde as duas bonecas foram colocadas na posição sentada, de modo que o participante ficou de frente para as bonecas. O experimentador ficou posicionado atrás do participante e realizava as perguntas presentes no roteiro semiestruturado. O participante foi instruído a imaginar que cada boneca era uma pessoa e que deveriam apontar para boneca que eles considerassem que possuía as características mencionadas. Foram perguntadas seis questões sobre adjetivos positivos (Alguma é Boa/Legal/Inteligente/Confiável/Forte/Bonita?) e seis adjetivos negativos antagônicos (Alguma é Má/Chata/Burra/Perigosa/Fraca/Feia?). O índice de viés

racial foi obtido a partir da diferença entre as frequências de atribuição de adjetivos positivos e negativos às bonecas branca e negra. O valor do índice aumentava à medida que adjetivos positivos eram atribuídos com maior frequência à boneca branca e adjetivos negativos à boneca negra, e diminuía quando ocorria o padrão inverso. O escore total variou de -12 a 12, sendo considerados indicativos de viés racial antinegro valores superiores a 4.

### **Intervenção**

Durante a intervenção, foram treinadas duas discriminações condicionais envolvendo um estímulo modelo unitário e dois estímulos de comparação, e quatro relações emergentes foram testadas por meio do procedimento de pareamento ao modelo simultâneo. O procedimento completo foi composto por três treinos principais (AB, BC e ABBC) e quatro testes correspondentes (BA, CB, AC e CA). Cada bloco de treino e teste foi composta por 16 tentativas, com exceção da sessão mista (ABBC), que consistiu em 32 tentativas (16 tentativas AB e 16 tentativas BC).

Nas sessões de treino, o estímulo modelo era apresentado na parte superior da tela, exigindo um clique do mouse. Após o clique, dois estímulos de comparação apareciam abaixo do modelo, posicionados no canto inferior direito e esquerdo da tela. Respostas corretas nas condições de treino produziam estrelas na tela por 2 s, enquanto respostas incorretas resultavam em uma tela preta por 1,5 s. A ordem de apresentação dos estímulos modelo e de comparação foi randomizada entre as tentativas. Durante os testes, não havia consequências diferenciais na tela, e o intervalo entre tentativas permaneceu constante em 1,5 s.

No treino AB, símbolos positivos (A1) ou negativos (A2) foram apresentados como estímulos modelo, e estímulos abstratos (B1 e B2) como comparações. As respostas corretas consistiam em escolher B1 na presença de A1 e B2 na presença de A2. Em seguida, foi conduzido o teste de simetria BA, no qual B1 e B2 foram apresentados como modelos e A1 e A2 como comparações, sem consequências diferenciais. No treino BC, os estímulos modelos foram B1 e B2, e os estímulos de comparação incluíram uma face negra (C1) e uma figura abstrata (C2). As respostas corretas consistiam em escolher C1 na presença de B1 e C2 na presença de B2. No teste de simetria CB, C1 ou C2 foram apresentados como modelos, e B1 e B2 como comparações, também sem consequências diferenciais. O critério de sucesso tanto para os treinos quanto para os testes de simetria foi estabelecido em 93% de acerto (15 de 16 tentativas).

No teste de transitividade AC, os estímulos modelo foram A1 ou A2, e os estímulos de comparação foram C1 e C2. No teste subsequente de equivalência CA, os estímulos modelo foram C1 e C2, e os estímulos de comparação foram A1 e A2. O critério para a formação de classes de equivalência foi estabelecido em 15 respostas corretas (de 16 tentativas) em ambos os testes AC e CA. Caso o participante não atingisse esse critério, era realizada uma sessão adicional de treino misto (ABBC), seguida de nova testagem de simetria. O treino misto e os testes foram repetidos até três vezes. Caso o critério não fosse alcançado, o procedimento era encerrado. Quinze dias após a intervenção, a manutenção das classes de equivalência foi verificada por meio dos testes AC e CA.

### **Validade Social**

Para avaliar a validade social da pesquisa sobre preconceito racial e da intervenção implementada, foi aplicado um questionário às cinco professoras responsáveis pelas turmas participantes do estudo. O questionário foi disponibilizado on-line 30 dias após a intervenção. A maioria das questões foi apresentada no formato de escala Likert de cinco pontos, e duas questões tinham formato aberto.

Inicialmente, buscou-se avaliar o quanto as professoras consideraram a pesquisa relevante, em uma escala que variava de 1 (nada relevante) a 5 (muito relevante). Em seguida, investigaram-se as percepções das professoras sobre a eficácia da intervenção na redução do preconceito racial. Foram incluídas perguntas sobre: (1) o impacto percebido do procedimento na frequência de insultos com conteúdo racial; (2) a forma como as crianças passaram a interagir com pessoas negras; (3) as escolhas das crianças por livros ou materiais envolvendo a cultura afro-brasileira; (4) a frequência de comentários positivos feitos pelas crianças sobre pessoas negras; e (5) a percepção geral de redução do viés antinegro. Nesses itens, a pontuação 1 indicava ausência de efeito percebido, enquanto a pontuação 5 indicava um efeito significativo.

Adicionalmente, foi solicitado que as professoras indicassem se identificaram ocorrências de preconceito racial, entre os participantes da pesquisa, antes e após a realização da pesquisa, com a possibilidade de relatar suas observações e experiências por meio de perguntas abertas.

## Resultados

### Matching-to-Sample

O treino AB foi repetido, em média, duas vezes até que o critério de aprendizagem fosse alcançado. Treze participantes atingiram a simetria BA na primeira tentativa, enquanto cinco necessitaram de uma repetição. O treino BC foi repetido, em média, 2,5 vezes. Quatro participantes atingiram a simetria CB na primeira tentativa, dez necessitaram de uma repetição e quatro precisaram de duas repetições. É relevante notar que o maior número de repetições necessárias ocorreu justamente no treino BC, que envolvia faces negras como estímulos, sugerindo que relações condicionais envolvendo esses estímulos foram adquiridas com maior dificuldade, possivelmente em função de redes relacionais pré-experimentais que tornaram essas contingências mais resistentes ao estabelecimento.

Todos os dezoito participantes atenderam aos critérios para a formação de classes de equivalência. Quatorze participantes atingiram o critério na primeira aplicação dos testes AC e CA. Quatro participantes não atingiram o critério nos testes AC e CA e passaram por uma repetição do treino misto. Desses, dois atingiram os critérios de equivalência imediatamente após o treino misto, e dois necessitaram de uma repetição adicional. Além disso, dezesseis participantes (88,88%) demonstraram manutenção das classes de equivalência quinze dias após a intervenção, com desempenho igual ou superior a 93,75% nos testes AC e CA. Os participantes P3 e P5 não mantiveram as classes de equivalência quinze dias após a intervenção.

## Teste Relacional

Todos os dezoito participantes apresentaram índices  $v$  superiores a 6 (critério de inclusão no estudo), indicando a presença de viés antinegro (ver Procedimento). A Tabela 2 apresenta, para cada participante, o índice  $v$  antes da intervenção (sombreado) e após a intervenção (não sombreado). A coluna mais à direita indica quantos pontos o índice  $v$  diminuiu na primeira sonda pós-intervenção em comparação com a última sonda pré-intervenção.

O principal efeito observado foi uma redução acentuada do índice  $v$  após a intervenção. Desconsiderando o participante P16, cuja redução foi de apenas 1 ponto após a intervenção, a diminuição de  $v$  para os outros dezessete participantes apresentou uma variação de 4 a 18 pontos, com mediana de 10 pontos. Para os participantes que tiveram duas ou três sondas pré-intervenção, o  $v$  permaneceu relativamente estável ao longo dessas sondas. O participante P16 foi novamente uma exceção, apresentando uma redução de cinco pontos entre a primeira e a terceira sonda pré-intervenção. Para os participantes que tiveram duas ou três sondas pós-intervenção, o  $v$  permaneceu abaixo do valor pré-intervenção, com exceção de P5 e P10 (novamente desconsiderando P16). Algumas outras crianças apresentaram aumento no  $v$ , mas com valores ainda inferiores aos observados no pré-intervenção. No follow-up, P5 e P13 retornaram ao mesmo padrão observado na linha de base. Por outro lado, P7, P9, P10 e P18 apresentaram índices variando de 0 a 4, enquanto os demais participantes exibiram índices superiores a 4 e inferiores a 8.

## Tabela 2

*Índice de Viés Racial Antinegro no Teste Relacional ao Longo dos Sondagens e do Follow-up*

	Sonda 1	Sonda 2	Sonda 3	Sonda 4	Follow-up
P1	12	2	4	2	8
P2	14	3	6	5	6
P3	16	2	3	6	8
P4	16	6	7	8	8
P5	16	8	8	16*	16*
P6	16	2	0	8	6
P7	10	10	-4	-3	2
P8	13	15	6	4	6
P9	16	16	3	4	4
P10	6	8	4	6	6
P11	16	16	6	4	4
P12	16	16	8	6	6
P13	16	16	16	6	16*
P14	14	15	14	8	6
P15	9	11	11	2	8
P16	13	9	8	7	8
P17	10	13	14	-4	7
P18	12	12	10	1	0

*Nota.* Asteriscos indicam retorno ao mesmo padrão antes da Intervenção

As mudanças pós-intervenção observadas no índice b decorreram principalmente da redução na frequência de relações entre faces negras e o símbolo negativo. Essa tendência é evidenciada nas matrizes apresentadas na Tabela 3, que mostram as frequências de escolha no teste relacional para cada um dos três estímulos de comparação (faces negras – C1, estímulos abstratos – C2 e faces brancas – C3), em função de cada estímulo modelo (símbolo positivo – A1; símbolo negativo – A2). As matrizes representam as médias das escolhas de 16 participantes, excluindo os outliers P5 e P16. A primeira coluna apresenta a média das escolhas nas sondas pré-intervenção; a segunda coluna apresenta as escolhas nos sondas pós-intervenção; e a terceira coluna apresenta a média das escolhas nas sondas de follow-up.

**Tabela 3**

*Percentual de escolhas de faces negras, faces brancas e estímulos abstratos em tentativas com estímulos modelo positivos e negativos.*

	Modelo	Antes da Intervenção Frequência (%)	Depois da Intervenção Frequência (%)	15 dias (Follow-up) Frequência (%)
FACES NEGRAS (C1)	A1 (+)	12.5%	37.5%	25%
	A2 (-)	87.5%	0%	0%
ABSTRATO (C2)	A1 (+)	12.5%	0%	0%
	A2 (-)	12.5%	100%	100%
FACES BRANCAS (C3)	A1 (+)	75%	62.5%	75%
	A2 (-)	0%	0%	0%

Nota. Oito tentativas foram apresentadas para cada Modelo (A1 and A2).

Na sonda pré-intervenção, os participantes predominantemente escolheram a face branca na presença do símbolo positivo e a face negra na presença do símbolo negativo. A mudança mais significativa observada na primeira sonda pós-intervenção foi o deslocamento das escolhas na presença do símbolo negativo, que passaram a se concentrar no estímulo abstrato. De forma notável, não houve seleções da face negra na presença do símbolo negativo, padrão que foi mantido no follow-up por 88,88% dos participantes.

Além disso, houve um aumento de três vezes nas escolhas da face negra na presença do símbolo positivo para a maioria dos participantes, com exceção de P5 e P14. No entanto, entre a sonda Sonda 4 e o follow-up, os participantes P1, P3, P4, P15 e P16 não selecionaram faces negras na presença do símbolo positivo em nenhuma tentativa. As faces brancas permaneceram sendo escolhidas com maior frequência na presença do símbolo positivo (em até 62,5% das tentativas) para a maioria dos participantes, com exceção de P7 nas sondas Sonda 3 e Sonda 4 e de P17 na sonda Sonda 4.

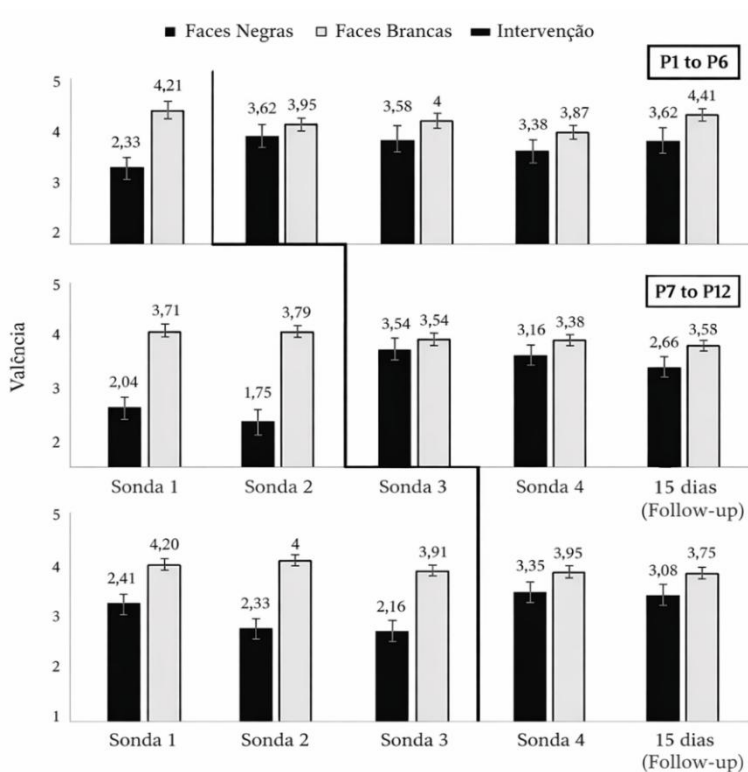
### **Escala de Valência Emocional (*Likert*)**

A Figura 2 apresenta as avaliações de faces negras e brancas na escala Likert. Para simplificar a apresentação, os resultados foram agrupados em médias para três grupos de participantes: aqueles que tiveram uma, duas ou três sondas pré-intervenção. As médias são apresentadas porque há pouca variabilidade entre participantes, como pode ser observado nos dados individuais disponíveis no material suplementar (ver Tabela 3, seção de Material Suplementar).

No pré-intervenção, as avaliações médias das faces brancas foram consideravelmente mais altas do que as das faces negras, sem sobreposição entre as distribuições. No pós-intervenção, houve pequenas reduções nas avaliações das faces brancas. Em contraste, as avaliações das faces negras apresentaram um aumento substancial, aproximando-se das avaliações das faces brancas, embora ainda ligeiramente inferiores. Os aumentos nas avaliações das faces negras mostraram-se relativamente estáveis, inclusive no follow-up.

#### **Figura 2.**

*Avaliações médias dos participantes antes e após a intervenção.*



Três participantes (P5, P7 e P13) retornaram ao padrão pré-intervenção, avaliando faces negras de forma negativa e faces brancas de forma positiva. Para os demais participantes, observou-se uma leve redução na valência atribuída às faces negras na sonda 4 e no follow-up. Ainda assim, essas avaliações permaneceram mais altas do que aquelas atribuídas no período pré-intervenção.

### Teste das Bonecas

Quinze participantes apresentaram consistentemente viés antinegro no teste das bonecas ( $v > 4$ ). A Tabela 3 apresenta o índice  $b$  no teste das bonecas antes da intervenção (sombreado) e após a intervenção (não sombreado). Nove participantes exibiram índice máximo de viés ( $v = 12$ ) nas sondas pré-intervenção. Dentre esses, P5, P11 e P14 não apresentaram mudanças após a intervenção, mantendo índices máximos ( $v = 12$ ). Por outro lado, quatro participantes (P1, P3, P4 e P10) demonstraram redução do índice de viés nas sondas pós-intervenção e

no follow-up. Dois participantes retornaram aos padrões pré-intervenção: P2 na sonda 4 e no follow-up, e P12 no follow-up.

Três participantes (P7, P8 e P9) apresentaram inicialmente índices de viés iguais a 8 nas sondas pré-intervenção. P8 reduziu o índice de viés em todas as sondas pós-intervenção e no follow-up, enquanto P9 reduziu o índice nas duas sondas após a intervenção, retornando a um padrão semelhante ao pré-intervenção no follow-up. Três participantes (P6, P15 e P16) apresentaram índices de viés iguais a 4, reduzindo o índice nas sondas pós-intervenção e no follow-up. O participante P13 alternou entre padrões de viés antinegro e pró-negro ao longo das sondas pré-intervenção; na sonda pós-intervenção, o índice foi 0, mas retornou a 12 no follow-up. Por fim, dois participantes (P17 e P18) não apresentaram viés antinegro nas sondas pré-intervenção do teste das bonecas e mantiveram esse padrão nas sondas pós-intervenção. A Tabela 4 apresenta índice de viés racial no teste da boneca ao longo das sondas e follow-up.

**Tabela 4**

*Índice de Viés Racial no Teste da Boneca ao Longo das Sondas e do Follow-up.*

	Sonda 1	Sonda 2	Sonda 3	Sonda 4	Follow-up
P1	12	0	-2	6	6
P2	12	-1	9	<b>12</b>	<b>12</b>
P3	12	4	4	0	4
P4	12	4	0	3	-1
P5	12	<b>12</b>	<b>12</b>	<b>12</b>	<b>12</b>
P6	4	0	2	2	2
P7	8	8	-2	0	0
P8	8	8	0	-1	-4
P9	8	8	1	2	<b>7</b>
P10	12	12	6	8	8
P11	12	12	<b>12</b>	<b>12</b>	<b>12</b>
P12	12	12	0	0	12
P13	12	-12	12	0	<b>12</b>
P14	12	12	12	<b>12</b>	<b>12</b>

P15	4	4	4	0	0
P16	4	4	4	-12	-4
P17	0	0	0	0	0
P18	0	0	0	0	0

Durante as sondas pré-intervenção, os participantes exibiram um padrão consistente: quando uma boneca era escolhida para um adjetivo positivo, a outra recebia sistematicamente o adjetivo negativo. Inicialmente, nove participantes selecionaram de forma consistente a boneca negra para todos os adjetivos negativos e a boneca branca para os adjetivos positivos. Seis participantes escolheram com maior frequência a boneca negra para adjetivos negativos (entre quatro e cinco tentativas) e a boneca branca para adjetivos positivos.

A Tabela 4 apresenta a frequência média (%) de relações entre bonecas negras e brancas e adjetivos positivos e negativos. Após a intervenção, a maioria dos participantes passou a exibir padrões de escolha mais flexíveis, por vezes selecionando ambas as bonecas ou nenhuma para determinados adjetivos. Diferentemente do teste relacional, a redução dos índices  $v$  no teste das bonecas esteve associada tanto à diminuição da frequência de relações entre a boneca negra e adjetivos negativos quanto ao aumento da frequência de relações com adjetivos positivos. Ainda assim, é importante notar que a redução das relações com adjetivos negativos foi mais pronunciada no teste relacional.

**Tabela 4**

*Percentual de escolhas de bonecas negras e brancas na presença de adjetivos positivos e negativos*

	Adjetivo	Antes da Intervenção	Depois da Intervenção	Follow-up
		Frequência (%)	Frequência (%)	Frequência (%)
Boneca Negra	Positivo	5.56%	65%	45.23%
	Negativo	94.44%	35%	53.57%
Boneca Branca	Positivo	94.44%	75.66%	88.10%
	Negativo	5.56%	19.6%	19.05%

*Nota.* Foram apresentados seis adjetivos positivos e seis adjetivos negativos.

### **Validade social**

Os resultados indicam que todos os cinco professores consideraram a pesquisa altamente relevante, com média de avaliação igual a 5, sugerindo que o tema do preconceito racial deve ser investigado e discutido no contexto escolar. Entretanto, em resposta às questões que avaliaram a efetividade da intervenção, os professores atribuíram escores entre 1 e 2 ( $M = 1,85$ ), indicando uma percepção de efeito baixo ou inexistente no contexto da sala de aula. Além disso, as crianças não mencionaram espontaneamente a pesquisa, e os professores não relataram a identificação de ocorrências de preconceito racial antes ou após a intervenção.

### **Discussão**

Este estudo aborda limitações identificadas em Passarelli et al. (2024) e investigou a generalização, manutenção e validade social dos efeitos decorrentes da formação de classes de equivalência entre faces negras e um símbolo positivo na redução do viés racial antinegro. Os achados de Passarelli et al. revelaram

discrepâncias entre as respostas dos participantes em medidas de seleção (por exemplo, teste relacional) e aquelas que avaliavam a generalização do treino (por exemplo, Escala Likert e Teste das Bonecas). Em resposta, foram realizadas adaptações para aumentar a precisão dessas medidas na identificação do viés antinegro. Além disso, a faixa etária foi ajustada para incluir crianças de 6 anos, diante de preocupações sobre possível ocultação de viés por participantes mais velhos (Krumpal, 2013).

Todos os participantes selecionados formaram classes de equivalência com sucesso. Esse resultado é relevante, pois estudos anteriores destacaram a dificuldade de formar classes de equivalência quando as relações treinadas entram em conflito com a história pré-experimental do indivíduo (e.g., de Carvalho & de Rose, 2014; Dixon et al., 2006; Haydu et al., 2015; Watt et al., 1991). O sucesso observado aqui pode ser atribuído a parâmetros facilitadores da formação de classes, como o protocolo simples para o complexo, no qual testes de simetria são realizados imediatamente após o treino de cada relação (Adams et al., 1993; Mizael et al., 2021). Contudo, observou-se variação no número de repetições dos blocos de treino, e participantes que necessitaram de mais repetições não mantiveram as classes quinze dias após a intervenção. Ainda assim, 88,88% dos participantes demonstraram manutenção das classes de equivalência quinze dias após a intervenção, replicando achados de estudos anteriores que utilizaram esse protocolo (Mizael et al., 2016, 2020; Passarelli et al., 2024).

Nas sondas pós-intervenção, todos os participantes apresentaram redução do viés antinegro, principalmente porque deixaram de relacionar faces negras ao símbolo negativo no teste relacional. Apenas os participantes que não mantiveram

as classes de equivalência voltaram a relacionar faces negras ao símbolo negativo. Houve um aumento relevante na frequência de relações entre faces negras e o símbolo positivo; entretanto, faces brancas continuaram sendo mais frequentemente relacionadas ao símbolo positivo após a intervenção. Isso sugere que a formação de classes foi altamente eficaz em extinguir a relação entre faces negras e o símbolo negativo, mas que as faces brancas permaneceram preferidas na presença do símbolo positivo (resultado semelhante ao de Mizael et al., 2016, 2020; Passarelli et al., 2024). Ainda assim, o desfecho é promissor, considerando a definição de preconceito racial como avaliações negativas de estímulos étnicos, como faces negras (Nogueira, 2006; Mizael & de Rose, 2017; Passarelli et al., 2023).

Na Escala Likert, os participantes inicialmente avaliaram faces negras de forma negativa e faces brancas de neutra a positiva, evidenciando uma clara diferença de valência antes da intervenção. Após a intervenção, observou-se um aumento da valência atribuída às faces negras, resultando em avaliações semelhantes para faces negras e brancas. Esses resultados replicam estudos anteriores (Mizael et al., 2016, 2020; Passarelli et al., 2024), com uma distinção importante: enquanto em estudos prévios a redução da diferença ocorreu principalmente por diminuição das avaliações das faces brancas, neste estudo a redução decorreu principalmente do aumento da valência das faces negras, sugerindo transferência de função do símbolo positivo após este tornar-se equivalente às faces negras (cf. Dougher et al., 1994; de Rose et al., 1988; Hughes et al., 2016; Ribeiro et al., 2020; Santos & de Rose, 2018, 2019). Embora faces brancas tenham permanecido mais relacionadas ao símbolo positivo, isso

não impediu o aumento da valência das faces negras na Escala Likert. Ainda assim, a valência das faces brancas permaneceu ligeiramente superior, o que pode explicar sua escolha mais frequente na presença do símbolo positivo no teste relacional, dado que estímulos com maior valência tendem a ser preferidos (Arntzen et al., 2016; Santos & de Rose, 2019; Smeets & Barnes-Holmes, 2003).

O aumento da valência das faces negras manteve-se relativamente estável nas sondas pós-intervenção e no follow-up para a maioria dos participantes. Contudo, na sonda 4 e no follow-up, três participantes (P6, P7, P13) retornaram aos padrões pré-intervenção na Escala Likert, enquanto os demais apresentaram apenas reduções leves. A transferência de função não é automática após o treino de equivalência (de Carvalho & de Rose, 2014; Ferro & Valero, 2008; Mizael et al., 2016, 2020; Passarelli et al., 2024) e, frequentemente, não se sustenta ao longo do tempo (Watters et al., 2023; Amd et al., 2018, 2019). Notavelmente, dois dos participantes que retornaram aos padrões pré-intervenção também não mantiveram as classes de equivalência. Isso evidencia a variabilidade individual na sustentabilidade dos efeitos e a importância de investigar variáveis que contribuam para a manutenção da aprendizagem. Dado que o viés racial é moldado e mantido pelos ambientes sociais das crianças, transformações induzidas por uma intervenção experimental breve podem não persistir após o retorno a esses ambientes, a menos que intervenções de manutenção sejam desenvolvidas.

No Teste das Bonecas, na medida inicial pré-intervenção, a maioria dos participantes (88,88%) apresentou índice de viés entre 4 e 12, indicando viés antinegro. Dezesesseis participantes selecionaram consistentemente a boneca

negra em 66,66% a 100% das tentativas com adjetivos negativos, em consonância com estudos clássicos (Clark & Clark, 1947; Camilo et al., 2020; Fernandes et al., 2008; França & Monteiro, 2002). Na sonda imediatamente pós-intervenção, 13 participantes reduziram o índice de viés, e 8 mantiveram essa redução até o follow-up. A redução decorreu tanto da diminuição de relações entre a boneca negra e adjetivos negativos quanto do aumento de relações com adjetivos positivos, indicando generalização avaliativa (Hughes et al., 2018; Fazio et al., 2004). Isso sugere que a equivalência entre faces negras e o símbolo positivo transformou funções avaliativas não apenas dos estímulos treinados, mas também de outros estímulos raciais, como as bonecas. Contudo, esse efeito mostrou-se limitado no seguimento: a transformação de funções avaliativas no Teste das Bonecas foi mantida em apenas metade dos participantes, em contraste com as medidas de classes de equivalência, nas quais o teste relacional demonstrou reversão em 16 dos 18 participantes, sugerindo que a generalização para estímulos novos é menos estável do que a manutenção das relações diretamente treinadas.

Antes da intervenção, observou-se um padrão dicotômico de respostas (e.g., se a boneca negra é “ruim”, a branca é “boa”). Após a intervenção, a maioria apresentou padrões mais flexíveis, selecionando ambas as bonecas ou nenhuma para certos adjetivos. Ainda assim, bonecas brancas permaneceram mais relacionadas a adjetivos positivos, e bonecas negras mantiveram alguma associação com adjetivos negativos, diferentemente do teste relacional, no qual faces negras deixaram de ser relacionadas ao símbolo negativo após a intervenção. Isso pode decorrer do fato de o teste relacional oferecer uma terceira

opção (estímulo abstrato), enquanto o teste das bonecas oferece apenas duas. Incluir um terceiro estímulo no teste das bonecas pode, em estudos futuros, reduzir ainda mais a escolha da boneca negra para adjetivos negativos.

Os resultados de validade social indicaram que todos os professores consideraram a pesquisa relevante e concordaram que o tema deve ser continuamente discutido. Contudo, relataram não perceber efeitos da intervenção no cotidiano das crianças ou em suas interações, o que pode indicar ausência de ocorrências explícitas ou baixa sensibilidade para detectar tais mudanças. Uma limitação do instrumento utilizado é que perguntas abertas e genéricas podem não orientar adequadamente a observação dos professores. Estudos futuros deveriam incluir questões mais específicas, descrevendo topografias de resposta concretas às quais os professores deveriam se atentar, como escolhas de parceiros em atividades, comentários sobre características físicas de colegas ou padrões de exclusão em situações de grupo, de modo a aumentar a sensibilidade da medida de validade social.

Em síntese, o estudo amplia achados anteriores (Mizael et al., 2016, 2021; Passarelli et al., 2024) com um desenho mais robusto, examinando efeitos do tempo e da generalização. Os resultados sugerem que a formação de classes de equivalência entre faces negras e um símbolo positivo promoveu transferência de valência e generalização para novos estímulos. Ademais, crianças de 6 anos parecem menos suscetíveis a respostas socialmente desejáveis (Levy & Hughes, 2009) e mais flexíveis à mudança (Belisle & Dixon, 2020; Lai et al., 2016).

### **Direções Futuras**

O preconceito racial é um fenômeno complexo e multifacetado (Passarelli et al., 2023; Matsuda et al., 2020; Parks & Kirby, 2022). Estudos futuros devem investigar medidas mais abrangentes para identificar preconceito racial e avaliar a generalização das intervenções. Para além da valência, é importante explorar outras propriedades dos estímulos raciais, como relações hierárquicas, preferências e tarefas de tomada de decisão.

O desenho adotado mostrou que a passagem do tempo pode deteriorar efeitos em algumas medidas (Escala Likert e Teste das Bonecas), ressaltando a necessidade de intervenções contínuas e follow-ups mais longos. Adaptar a intervenção a contextos naturalísticos (jogos, narrativas) pode favorecer a emergência de relações positivas e aumentar a validade social. Avaliar generalização em medidas mais próximas da vida real (e.g., escolha de parceiros para tarefas) também é essencial.

Por fim, embora a intervenção tenha sido eficaz com crianças de 6 anos, pode não ser tão efetiva com crianças mais velhas ou adultos, que possuem histórias mais extensas de responder relacional e menor flexibilidade. Dado que o viés antinegro se associa à inflexibilidade relacional (Flynn et al., 2018; Levin et al., 2016; Florez et al., 2019), intervenções que aumentem a flexibilidade podem ser promissoras para esses grupos (Matsuda et al., 2020; Kishita et al., 2014; Shea et al., 2023; Passarelli et al., 2024).

## Material Suplementar

**Tabela 5**

*Valência atribuída às faces brancas e negras durante às sondas e follow-up*

	Sonda 1		Sonda 2		Sonda 3		Sonda 4		15 dias (Follow-up)	
	Faces negras	Faces Brancas	Faces negras	Faces Brancas	Faces negras	Faces Brancas	Faces negras	Faces Brancas	Faces negras	Faces Brancas
P1	1.5	4	3.25	3.25	4.75	5	3.25	3.25	4.25	5
P2	2.25	3.75	3.75	3.75	4	4.25	4	4.25	4	4.25
P3	2.75	5	5	5	4	4	4.75	5	4	5
P4	2.5	4.5	3.75	3.5	2.75	2.75	2.5	2.5	4	4
<b>P5</b>	<b>2.75</b>	<b>4.25</b>	<b>3</b>	<b>4.25</b>	<b>3</b>	<b>4.25</b>	<b>2.75</b>	<b>4.25</b>	<b>2.5</b>	<b>4.25</b>
P6	2.25	3.75	3	4	3	3.75	3	4	3	4
<b>P7</b>	<b>3</b>	<b>4</b>	<b>2.5</b>	<b>4.5</b>	<b>4</b>	<b>4.25</b>	<b>3.25</b>	<b>4</b>	<b>2.5</b>	<b>3.75</b>
P8	1	4	1	3.5	3.25	3.5	2	2.25	2.75	3.5
P9	2	3.5	1	3	4	4	5	4.5	3.5	3.25
P10	2.5	4.25	2	4.5	2.75	3.5	3.25	4	3	4.25
P11	2	3.5	2.25	4.25	3	3.5	3.5	4.25	2.75	4.25
P12	2	3.5	1.5	3	1.5	3.5	2.5	3.5	2.5	3
<b>P13</b>	<b>1.75</b>	<b>3</b>	<b>1.75</b>	<b>3</b>	<b>2.25</b>	<b>2.5</b>	<b>2</b>	<b>2.5</b>	<b>1.5</b>	<b>2</b>
P14	2.5	4.75	2	5	2.75	4.25	5	5	4.5	4.5
P15	2.25	4.5	2.25	3.75	2.25	4	2.75	4.25	2.75	4
P16	2.75	4.25	2.25	4.5	1.5	4	2.75	4	2.25	4.25
P17	2.5	4	2.75	3.75	2	4	2.75	3.5	2.75	3.25
P18	2.5	4.25	3.25	4	3	3.75	3.75	3.5	3.75	3.5

*Nota.* As colunas sombreadas indicam a fase pré-intervenção. Os participantes que retornaram ao padrão pré-intervenção estão destacados em negrito.

-----

## Comentário

### **Da Eficácia Experimental à Validade Social: Alcances e Limites**

Esse estudo soma-se a outros presentes na literatura para responder uma questão importante desta tese: a formação de classes de equivalência entre faces negras e estímulos positivos é suficiente para reduzir avaliações negativas dessas faces em contextos controlados? A resposta, agora replicada com diferentes delineamentos e faixas etárias, é afirmativa. Sempre que as classes foram formadas e mantidas, observou-se a eliminação da relação entre faces negras e estímulos negativos no teste relacional e a redução das diferenças de valência entre faces brancas e negras em medidas avaliativas diretas. Ao mesmo tempo, esses dados confirmam que a simples exposição repetida às medidas, sem intervenção, não produz mudanças sistemáticas no viés racial.

Mais uma vez, participantes que não formaram classes de equivalência, ou que não as mantiveram ao longo do tempo, continuaram a apresentar padrões de viés racial semelhantes aos observados antes da intervenção. Essa convergência negativa evidencia a relação funcional direta entre o procedimento de treino relacional e a mudança observada: quando a reorganização das relações condicionais não ocorreu, a mudança avaliativa também não ocorreu. Além disso, mesmo em crianças pequenas, para as quais se poderia supor maior flexibilidade relacional, padrões de rigidez e resistência à mudança estiveram presentes, revelando variabilidade individual no impacto do procedimento.

Os efeitos foram, ainda assim, mais amplos do que aqueles observados em estudos com crianças de oito a dez anos. A maioria dos participantes apresentou maior convergência entre medidas explícitas antes da intervenção e, após o procedimento, evidência de generalização para o Teste da Boneca e maior flexibilidade relacional. Alguns

demonstraram redução do viés racial com escolhas mais equilibradas entre as bonecas; outros apresentaram inversão do padrão inicial em algumas sondas, sugerindo enfraquecimento do controle dicotômico previamente estabelecido. Ainda assim, a generalização não foi observada para todos os participantes, reforçando que a transformação de funções derivadas pode alcançar medidas socialmente mais complexas, mas não de forma uniforme.

Os dados de validade social acrescentaram uma dimensão crítica. Embora os professores tenham considerado o tema altamente relevante, relataram poucos efeitos observáveis no cotidiano escolar. Esse descompasso entre eficácia experimental e ausência de mudanças perceptíveis no ambiente natural evidencia um limite do procedimento tal como aplicado. A eficácia experimental, por si só, não é suficiente quando o objetivo envolve práticas educacionais e contextos sociais complexos. A questão que se impõe deixa de ser apenas se o procedimento funciona, e passa a ser onde, como e para quem ele funciona de maneira relevante.

Essas questões orientaram as decisões metodológicas do estudo seguinte. O retorno ao paradigma de equivalência de estímulos não representa uma ruptura, mas um reencontro informado pelos dados acumulados. Desta vez, o objetivo central não foi demonstrar que o procedimento produz mudanças avaliativas, mas examinar a generalidade desses efeitos sob variações contextuais sistemáticas e com maior aproximação ao ambiente natural. Para isso, três mudanças foram introduzidas. Primeiro, os estímulos avaliativos foram substituídos: em lugar de símbolos positivos e negativos arbitrários, foram utilizadas palavras positivas e negativas frequentemente envolvidas em estereótipos raciais no cotidiano social, o que aumenta a validade ecológica do treino e aproxima o procedimento das contingências culturais nas quais o viés racial é adquirido e mantido. Segundo, o Teste da Boneca foi

reformulado para incluir perguntas sobre atividade física, desempenho acadêmico e amizade, ampliando o escopo para além da dimensão avaliativa e aproximando as medidas de contextos de discriminação socialmente relevantes. Terceiro, o estudo foi conduzido em uma escola com histórico consolidado de ações antirracistas e com inserção institucional mais ampla da pesquisadora, que incluiu oficinas, atividades educativas e articulação contínua com professores e gestores, criando condições favoráveis tanto para a coleta quanto para a avaliação da validade social. Essas mudanças não alteram a lógica do procedimento, mas ampliam seu escopo e sua relevância para contextos educacionais reais.

\*\*\*\*\*

## **Estudo 6 - Efeitos da Formação de Classes de Equivalência na Redução do Preconceito Racial em Crianças II: Evidências de Generalização, Manutenção e Validade Social<sup>11</sup>**

O preconceito racial pode ser compreendido como um conjunto de respostas avaliativas negativas aprendidas e mantidas por práticas culturais e redes simbólicas compartilhadas. Estudos anteriores demonstraram que a formação de classes de equivalência entre faces negras e estímulos positivos, por meio de procedimentos de *matching-to-sample*, é eficaz para reduzir avaliações negativas, com evidências de transferência de valência e generalização. O presente estudo teve como objetivo avaliar a generalidade desses efeitos utilizando palavras estereotípicas de valência positiva e negativa, bem como examinar seus efeitos sobre diferentes medidas de viés racial. Inicialmente, 100 crianças brasileiras (M = 6 anos) foram avaliadas por meio de tarefas relacionais, avaliação afetiva de valência, preferência e discriminação em atividades sociais. Quinze crianças com viés racial antinegro estável em pelo menos três medidas foram selecionadas para a intervenção. O procedimento empregou um protocolo simples-para-complexo de *matching-to-sample* envolvendo palavras estereotípicas, estímulos abstratos e faces negras. Treze participantes (86,66%) formaram classes de equivalência, com manutenção até 45 dias. Observou-se evidência de generalização para a Escala SAM (13/15; 86,66%) e para o Teste da Boneca (8/15; 53%), ainda que com variação ao longo das sondas. Os resultados indicam potencial translacional do procedimento para intervenções educativas na infância.

---

<sup>11</sup> Passarelli, D.A., Mizaél, T.M., Roche, B., De Rose, J. C. (2026). Efeitos da Formação de Classes de Equivalência na Redução do Preconceito Racial em Crianças II: Evidências de Generalização, Manutenção e Validade Social. [Manuscrito em Processo de Preparação para ser submetido em periódico científico]

## Introdução

O preconceito racial pode ser compreendido como um conjunto de respostas avaliativas negativas direcionadas a indivíduos ou grupos com base em características étnico-raciais, sendo historicamente moldado e mantido por práticas culturais, contingências sociais e redes simbólicas compartilhadas (Nogueira, 2007; Mizael & de Rose, 2017; Passarelli et al., 2023). Na perspectiva da análise do comportamento, tais respostas podem ser investigadas a partir de relações estabelecidas entre estímulos raciais e funções avaliativas positivas ou negativas, bem como das condições sob as quais essas funções são adquiridas, mantidas e transformadas (Mizael et al., 2016, 2020; Passarelli, Roche et al., 2024).

Uma linha consistente de pesquisas empíricas tem demonstrado que o paradigma de equivalência de estímulos constitui uma ferramenta robusta e eficaz para investigar e modificar respostas avaliativas e preferências (Barnes-Holmes et al., 2000; Bortoloti & de Rose, 2009; Mizael, dos Santos, et al., 2016; Grey & Barnes 1996; dos Santos & de Rose, 2019), o que também tem ocorrido com estímulos raciais (Mizael, de Almeida, et al., 2016, 2020; Passarelli, Roche et al., 2024; Passarelli et al., 2026). A formação de classes de equivalência permite que funções previamente atribuídas a um estímulo sejam transferidas para outros membros da classe, produzindo mudanças avaliativas mesmo na ausência de treino direto (de Rose et al., 1988; Dougher et al., 1994; Hughes et al., 2018). Em estudos com crianças, procedimentos baseados em equivalência têm sido

utilizados para atenuar a negatividade das avaliações dirigidas a faces negras (Mizael et al., 2016, 2020), com evidências de generalização para novas medidas e manutenção desses efeitos mesmo quinze dias após a intervenção (Passarelli, Roche, et al., 2024).

No entanto, grande parte desses estudos utilizou como estímulos afetivos símbolos positivos e negativos arbitrários como elementos centrais no treino de equivalência para promover a transferência de propriedades avaliativas. Embora tais estímulos sejam eficazes para demonstrar experimentalmente a transferência de função, eles tendem a apresentar história pré-experimental limitada quando comparados a palavras que fazem parte do repertório verbal cotidiano das crianças. Palavras estereotípicas associadas a atributos positivos ou negativos integram redes simbólicas amplamente compartilhadas culturalmente e exercem controle sobre o comportamento em múltiplos contextos sociais, incluindo interações interpessoais, narrativas, jogos, práticas educativas e atividades lúdicas (Passarelli et al., 2023).

Evidências empíricas indicam que estímulos com funções pré-experimentais estabelecidas podem interferir na formação e reorganização de classes de equivalência, sobretudo quando as relações treinadas entram em conflito com a história verbal prévia dos participantes (Haydu et al., 2015, 2019; Moxon et al., 1993; Watt et al., 1991). Do ponto de vista conceitual, esses efeitos refletem o conflito funcional entre repertórios relacionais previamente estabelecidos e relações simbólicas recentemente aprendidas, nas quais relações historicamente consolidadas tendem a exercer controle preferencial sobre o responder, dificultando ou retardando a aquisição de relações inconsistentes com

a história cultural do indivíduo (ver Mizael et al., 2016, 2020; ver também Passarelli et al., 2020; Passarelli, Roche et al., 2024) Assim, dificuldades na formação de classes não devem ser interpretadas como falhas experimentais, mas como expressões do peso funcional do significado e das redes simbólicas previamente selecionadas, conforme discutido no âmbito do paradigma da equivalência de estímulos (de Rose & Bortoloti, 2007; Bortoloti & de Rose, 2014). Nesse sentido, investigar a formação e reorganização de classes envolvendo palavras estereotípicas aproxima o procedimento experimental das condições sob as quais o preconceito racial é aprendido e mantido no cotidiano infantil, ampliando sua validade ecológica.

Além disso, avaliar a generalidade de procedimentos baseados em equivalência utilizando palavras estereotípicas representa um avanço relevante, pois permite examinar se a formação de classes de equivalência é robusta o suficiente para reorganizar funções avaliativas quando os estímulos envolvidos possuem história verbal prévia mais extensa (Haydu et al., 2015; Haydu et al., 2015, 2019; Moxon et al., 1993; Watt et al., 1991; Belisle & Dixon, 2020). O uso de palavras aproxima o procedimento experimental de contextos nos quais o viés racial é usualmente aprendido e mantido, contribuindo para a análise de sua relevância translacional.

Outro aspecto central para a avaliação da generalidade desses procedimentos refere-se à ampliação do conjunto de medidas utilizadas para identificar mudanças no repertório relacionado ao viés racial. Medidas baseadas exclusivamente em avaliações de valência capturam apenas uma dimensão do controle exercido por estímulos raciais, podendo subestimar a complexidade

funcional desse repertório (Passarelli et al., no prelo). A inclusão de tarefas que envolvem escolhas relacionais, preferências e discriminação em atividades sociais (Clark & Clark, 1947; Camilo et al., 2020; Passarelli, Roche et al., 2024), bem como medidas de identificação racial (Clark & Clark, 1947) e avaliações de diferentes dimensões afetivas (Bradley & Lang, 1994; Lang et al., 1999), permite examinar como diferentes contextos de resposta organizam padrões distintos de comportamento sob controle de estímulos raciais. Essa abordagem multimodal possibilita uma análise mais abrangente dos efeitos da intervenção, ao avaliar a consistência, a variabilidade e a generalização dos efeitos observados entre tarefas com diferentes demandas comportamentais e graus de explicitação (Campbell & Fiske, 1959; Nosek & Smyth, 2007).

Adicionalmente, a inclusão de indicadores de validade social tem sido apontada como um componente importante na avaliação de intervenções voltadas à modificação de repertórios socialmente relevantes, especialmente quando se trata de fenômenos complexos e culturalmente sensíveis, como o preconceito racial (Matsuda et al., 2020; Parks & Kirby, 2022). A percepção de observadores sobre possíveis mudanças no comportamento das crianças pode fornecer informações complementares sobre a aceitabilidade e a relevância social dos procedimentos empregados.

Diante dessas considerações, o presente estudo teve como objetivo avaliar a generalidade de um procedimento baseado na formação de classes de equivalência utilizando palavras estereotípicas de valência positiva e negativa, examinando seus efeitos sobre diferentes respostas relacionadas ao viés racial em crianças. Especificamente, investigou-se a formação de classes de

equivalência e a generalização de seus efeitos para um conjunto ampliado de medidas, incluindo tarefas de escolha forçada envolvendo faces negras e brancas, testes relacionais, preferências e discriminação social no Teste da Boneca adaptado, identificação racial, avaliações afetivas por meio da Escala SAM (valência) e indicadores de validade social obtidos junto à equipe da escola.

## **Método**

### **Participantes**

Cem crianças brasileiras (M = 6 anos) de uma escola estadual do interior do estado de São Paulo foram inicialmente avaliadas por meio do Teste Relacional, Teste de Preferência Relacional, Escala *Self-Assessment Manikin* (SAM) e Teste da Boneca modificado (ver seção *Instrumentos*).

Os critérios de seleção para participação neste estudo buscaram garantir que apenas participantes que apresentassem um padrão consistente de viés racial desfavorável em relação a faces negras fossem incluídos. Para isso, exigiu-se que os participantes demonstrassem esse viés em pelo menos três das quatro medidas avaliativas utilizadas no estudo, as quais envolviam diferentes tipos de tarefas relacionais, preferências sociais e avaliações afetivas (descritas em detalhe na seção *Instrumentos*). Além disso, foi necessário que esse padrão de viés se mantivesse estável ao longo de sucessivas avaliações iniciais, realizadas antes da intervenção. Em geral, exigiu-se a manutenção do viés ao longo de três sondagens consecutivas. Nos casos em que foi observado um enfraquecimento gradual do viés durante essa fase inicial, a inclusão foi condicionada à presença de pelo menos duas linhas de base estáveis, conforme os critérios de estabilidade descritos posteriormente na seção *Procedimento*.

Quinze crianças atenderam aos critérios de inclusão supracitados (M = 6 anos; 10 do sexo feminino). A classificação étnico-racial dos participantes foi realizada pelos pais ou responsáveis legais no momento do preenchimento do Termo de Consentimento Livre e Esclarecido, sendo 7 participantes (46,67%) classificados como brancos e 8 (53,33%) como pardos.

### **Procedimento**

Foi utilizado um *Delineamento de Sondagens Múltiplas* não concorrentes (Cooper et al., 2020) entre participantes. Cada criança participou de três a quatro sondagens de linha de base, com dois dias de espaçamento entre as sondagens, conduzidas antes da intervenção. As sondagens de linha de base (pré-intervenção) objetivaram verificar a estabilidade do viés racial ao longo do tempo.

Em cada sondagem, foram aplicadas as quatro medidas avaliativas (ver Instrumentos e Estímulos), sempre na mesma ordem, a fim de padronizar o contexto experimental. As medidas foram conduzidas individualmente, em ambiente silencioso da escola, com a criança posicionada em frente ao computador ou à mesa experimental, conforme a natureza da tarefa.

### **Sondagens**

As sondagens avaliativas consistiram em testes relacionais computadorizados, uma medida pictórica de valência afetiva e uma tarefa explícita de preferência social. Esses procedimentos foram aplicados repetidamente ao longo do estudo com o objetivo de avaliar a estabilidade inicial do viés racial, selecionar participantes elegíveis e examinar mudanças após a intervenção.

*Teste Relacional com Três possibilidades de Escolha.* Teste conduzido por meio de um procedimento de *matching-to-sample* (MTS), conforme descrito por

Mizael et al. (2016) e utilizado em estudos subsequentes (Mizael et al., 2020; Passarelli, Roche et al., 2024). O teste avaliou o padrão de responder relacional entre adjetivos positivos ou negativos e estímulos raciais ou neutro. Em cada tentativa, um estímulo-modelo avaliativo A1 (positivo: *bonito, calmo, inteligente ou bom*) ou A2 (negativo: *feio, agressivo, burro ou mau*) era apresentado isoladamente na tela. A apresentação dos estímulos de comparação era contingente à resposta do participante ao estímulo-modelo, que consistia em um clique sobre o modelo apresentado. Imediatamente após essa resposta, eram apresentados simultaneamente três estímulos de comparação: uma face negra (C1), um estímulo abstrato (C2) e uma face branca (C3), dispostos em posições randomizadas na tela.

Os participantes eram instruídos a selecionar o estímulo de comparação que considerassem correspondente ao modelo. O teste foi composto por 16 tentativas, sendo oito com o símbolo positivo e oito com o símbolo negativo, apresentadas em ordem

O viés racial foi quantificado por meio de um índice de viés racial ( $b$ ), calculado a partir da frequência de escolhas de faces brancas e negras na presença dos adjetivos positivos e negativos, conforme a fórmula proposta por Mizael et al. (2016). Os valores variam de +16 a -16, com valores positivos indicando maior viés antinegro (ou seja, face negras negativas e brancas positivas) e valores negativos menor viés antinegro (sendo brancas negativas e negras positivas). Foram selecionados participantes com índice de viés maior 6 ao longo das sondas.

*Teste de Preferência Relacional com Duas Escolhas.* Este teste seguiu o mesmo formato geral do teste relacional descrito anteriormente, diferindo apenas pelo número de estímulos de comparação apresentados. Em vez de três alternativas, foram apresentadas apenas duas comparações, uma face negra e uma face branca, com o objetivo de examinar escolhas em um contexto dicotômico (cf. Passarelli, Roche et al., 2024). Todos os demais parâmetros do procedimento, incluindo apresentação dos estímulos-modelo, número de tentativas, balanceamento e ausência de consequências diferenciais, foram idênticos aos do teste anterior.

A criança era instruída a observar o adjetivo apresentado na tela e a escolher qual das duas faces (uma branca e uma negra) melhor combinava com aquele adjetivo. Essa tarefa selecionava um contexto de escolha relacional sob restrição de alternativas, avaliando como adjetivos avaliativos positivos e negativos controlavam a seleção entre faces raciais quando não havia opção neutra disponível. A preferência racial foi quantificada por meio do mesmo índice de viés ( $b$ ) utilizado no Teste Relacional, calculado a partir da frequência de escolhas de faces brancas e negras diante dos adjetivos positivos. Valores mais elevados de  $b$  indicam maior preferência relativa por faces brancas na presença de adjetivos positivos. Foram selecionados participantes com índice  $b$  maior que 6 em todas as sondas.

*Escala Self-Assessment Manikin (SAM).* Apenas a dimensão valência da escala foi empregada, uma vez que as crianças demonstraram não compreender adequadamente as dimensões de excitação (arousal) e dominância, apesar das instruções fornecidas. As imagens de rostos foram obtidas do banco

*FaceResearch* (faceresearch.org) e consistiram em 16 fotografias com expressão facial neutra, distribuídas igualmente quanto a gênero e raça: quatro faces femininas negras, quatro faces femininas brancas, quatro faces masculinas negras e quatro faces masculinas brancas. As faces foram apresentadas individualmente no centro da tela, de forma balanceada quanto ao gênero e à raça.

A escala utilizada consistiu em uma escala Likert de cinco pontos, representada por figuras pictóricas que expressavam diferentes níveis de valência afetiva. Os pontos 1 e 2 indicavam valência negativa, por meio de figuras com expressão facial triste; o ponto 3 indicava valência neutra, com expressão neutra; e os pontos 4 e 5 indicavam valência positiva, representada por figuras com expressão facial feliz. A escala foi apresentada imediatamente abaixo do estímulo-alvo avaliado.

A criança era instruída a observar a face apresentada no centro da tela e a escolher a figura que melhor representasse o quanto ela gostava ou não gostava daquele rosto. Foi explicado que as figuras pictóricas variavam de uma expressão mais triste para uma mais feliz. Caso a criança gostasse muito da face apresentada, deveria selecionar a figura pictórica 5, caracterizada por maior curvatura de sorriso; caso apenas gostasse, deveria selecionar a figura 4, com menor curvatura de sorriso; caso avaliasse a face como neutra, sem gostar ou desgostar, deveria selecionar a figura 3; caso desgostasse, deveria selecionar a figura 2, representada por expressão triste com curvatura descendente; e, caso desgostasse muito, deveria selecionar a figura 1, correspondente à expressão mais negativa, com maior curvatura descendente e expressão muito triste. Foi considerado indicativo de viés racial antinegro o padrão no qual a criança atribuía

avaliações negativas (1 ou 2 pontos) às faces negras e avaliações superiores às faces brancas.

*Teste da Boneca adaptado.* Utilizado como medida explícita de preferência e discriminação racial, inspirado no paradigma clássico de Clark e Clark (1947), com adaptações metodológicas. O conjunto de estímulos foi composto por cinco desenhos de bonecas fisicamente idênticas, que diferiam exclusivamente quanto ao tom de pele, organizadas em um contínuo do mais claro ao mais escuro. As bonecas 1 e 2 representavam tonalidades brancas; a boneca 3, tonalidade parda; e as bonecas 4 e 5, tonalidades negras, sendo a 4 de tom mais claro e a 5 de tom mais escuro (ver Figura 2, Experimento 2).

As crianças responderam a perguntas padronizadas de preferência em diferentes contextos de atividades sociais (e.g., “Escolheria alguma para ser sua amiga?”, “Escolheria alguma para fazer um trabalho para nota?”, “Escolheria alguma para o time da Educação Física?”), podendo selecionar livremente nenhuma, uma ou mais bonecas em cada pergunta. Adicionalmente, para investigar respostas de discriminação social, foi perguntado se alguma das bonecas não seria escolhida para cada atividade. Incluiu-se também a pergunta “Desconfia de que alguma roubaria suas coisas?”, com o objetivo de avaliar estereótipos negativos relacionados à criminalidade atribuídos diferencialmente em função do tom de pele.

Cinco imagens de bonecas, com gradações de pele diferente, eram apresentadas no computador. A criança estava posicionada em frente ao computador e o experimentador sentado ao lado da criança. A criança era instruída a imaginar que as bonecas representavam crianças novas que haviam

entrado na escola e deveria responder às perguntas sobre esses alunos novos feitas pelo experimentador apontando qual seria o selecionado. As instruções enfatizavam que a criança poderia escolher nenhuma, uma ou mais bonecas, conforme achasse adequado em cada pergunta.

Foi considerado viés racial quando participantes selecionavam boneca branca para atividades sociais (preferências) e não selecionavam para atividade criminosa; e quando não selecionavam as negras para atividades sociais e escolhiam a negra para atividade criminosa (discriminação). A seleção ou não seleção da variável categórica (boneca) foi registrada em cada sonda.

### **Intervenção**

A intervenção empregou o mesmo protocolo de formação de classes de equivalência utilizado em estudos anteriores dessa linha de pesquisa (Mizael et al., 2016, 2020; Passarelli, Roche et al., 2024), baseado em um procedimento de *matching-to-sample* simultâneo (SMTS) organizado segundo o arranjo simples-para-complexo. Os estímulos consistiram em A1 (bonito, calmo, inteligente e bom) e A2 (feio, agressivo, burro e mau), respectivamente, B1 e B2 (estímulos abstratos) e classes de estímulos C1 e C2. A classe C1 consistiu em quatro faces negras (dois homens e duas mulheres), e C2 foi composta por um estímulo abstrato.

Foram treinadas duas discriminações condicionais, cada uma envolvendo dois estímulo-modelo e dois estímulos de comparação, e após uma linha de base cumulativa misturando as duas discriminações condicionais, o teste das relações emergentes. O protocolo completo foi composto por três treinos principais (AB, BC e ABBC) e quatro testes correspondentes (BA, CB, AC e CA). Cada bloco de

treino e teste incluiu 16 tentativas, com exceção do bloco misto ABBC, que consistiu em 32 tentativas (16 tentativas AB e 16 tentativas BC).

Nas sessões de treino, o estímulo-modelo era apresentado na parte superior da tela, seguido pela apresentação de dois estímulos de comparação posicionados abaixo, de modo que as posições foram randomizadas. Respostas corretas produziram consequências visuais breves (estrelas na tela), enquanto respostas incorretas resultaram em uma tela neutra (cinza). Durante os testes, não houve consequências diferenciais para as respostas, e o intervalo entre tentativas permaneceu constante.

No treino AB, palavras de valência positiva (A1) ou negativa (A2) foram apresentadas como estímulos-modelo, e estímulos abstratos (B1 e B2) como comparações. No treino BC, os estímulos-modelo foram os estímulos abstratos (B1 e B2), e os estímulos de comparação incluíram uma face negra (C1) e um estímulo abstrato (C2). Os testes de simetria (BA e CB) foram conduzidos sem feedback, imediatamente após cada treino correspondente, conforme o protocolo simples-para-complexo.

O critério de desempenho tanto nos treinos quanto nos testes de simetria foi estabelecido em 93% de acerto (15 de 16 tentativas). O treino misto (ABBC) era implementado apenas após o participante atingir critério nos treinos AB e BC e em suas respectivas simetrias. A formação de classes de equivalência foi avaliada por meio dos testes de transitividade (AC) e simetria da transitividade (CA), adotando-se o mesmo critério de desempenho (15 acertos).

Caso o critério não fosse alcançado nos testes de equivalência, o treino misto era reaplicado, seguido de nova testagem. Esse procedimento podia ser

repetido até três vezes. Caso o critério não fosse atingido após a terceira repetição, o procedimento de intervenção era encerrado para aquele participante e a sonda com as medidas avaliativas era realizada.

Cabe destacar que houve de três a quatro sondas (a depender do número de sondas da linha de base) com medidas avaliativas após a intervenção. As sondas foram aplicadas imediatamente após a intervenção, dois e quatro dias após a intervenção e para alguns participantes seis dias após a intervenção.

A manutenção das classes de equivalência foi avaliada quinze, trinta e quarenta e cinco dias após a intervenção por meio da reaplicação dos testes AC e CA. Juntamente com o teste de manutenção, foram aplicadas as sondas com as quatro medidas avaliativas.

### **Validade Social**

A validade social do estudo e da intervenção foi avaliada conforme os critérios propostos por Wolf (1978), contemplando a relevância social dos objetivos, a aceitabilidade dos procedimentos e a importância prática dos efeitos observados. Para isso, foi aplicado um questionário a dois professores responsáveis pelas turmas participantes, ao coordenador pedagógico e ao diretor da escola (N = 4).

O questionário foi disponibilizado em formato on-line 30 dias após a conclusão da intervenção. A maior parte das questões foi apresentada em escala Likert de cinco pontos, e duas questões abertas permitiram o relato qualitativo de observações feitas no contexto escolar.

Inicialmente, avaliou-se a relevância percebida da pesquisa, por meio de uma escala que variava de 1 (*nada relevante*), 2 (*pouco relevante*), 3 (*relevante*) 4

(*muito relevante*) a 5 (*extremamente relevante*). Em seguida, investigou-se a eficácia e efetividade percebida da intervenção na mesma escala de 1 (*nada*) a 5 (*extremamente*), com itens que abordaram: (1) mudanças na ocorrência de comportamentos verbais com conteúdo racial negativo; (2) alterações na forma como as crianças passaram a interagir com colegas e adultos negros; e (3) a percepção geral de redução do viés racial no ambiente escolar. Os respondentes podiam comentar suas avaliações com exemplos.

Adicionalmente, os respondentes foram solicitados a indicar se haviam observado ocorrências de preconceito racial antes e após a intervenção, com espaço para descrição livre de exemplos ou situações relevantes. As perguntas abertas tiveram como objetivo captar evidências de generalização dos efeitos da intervenção para o contexto natural da escola, bem como possíveis mudanças socialmente relevantes não diretamente mensuradas pelas tarefas experimentais.

## **Resultados**

### **Desempenho na Intervenção**

Treze (86.66%) dos participantes demonstraram formação de classes de equivalência, com 100% de acurácia nos testes de transitividade (AC) e simetria da transitividade (CA), na primeira oportunidade do teste. Para esses participantes, o critério de aprendizagem (15 tentativas consecutivas) no treino AB foi atingido com a média de 23.45 tentativas e no treino BC foi atingido em 28.76. Todos os participantes passaram nos testes de simetria na primeira oportunidade do teste. Além disso, todos demonstraram manutenção das classes de equivalência 15, 30 e 45 dias após a intervenção.

Dois participantes (P14 e P15) necessitaram de duas novas exposições ao treino ABBC e foram expostos a duas novas oportunidades de teste AC e CA, mas não formaram classes de equivalência. Após a terceira tentativa no teste AC e CA, foi considerado que não houve formação de classes de equivalência. Para esses participantes, a média de tentativas para atingir critério no treino AB foi similar aos outros ( $M= 24.75$ ). No entanto, no treino BC o critério de aprendizagem foi atingido apenas na trigésima quinta tentativa para P14 e quadragésima sexta para P15, considerando a primeira oportunidade de treino. Já no teste de simetria, participantes atingiram critério na primeira oportunidade

Como as classes não foram formadas na primeira oportunidade, foi empregado treino misto (ABBC), seguido do teste de simetria (BA e CB) e equivalência (AC e CA). Cabe destacar que P14 apresentou apenas 3 erros no treino misto até atingir critério e enquanto P15 apresentou 10 erros até atingir critério. Ambos apresentaram simetria na primeira oportunidade, mas não demonstraram equivalência em nenhuma das duas oportunidades adicionais (0 acertos).

### **Teste Relacional com Três Estímulos Comparações**

Cinco participantes (P1–P5) apresentaram viés racial máximo ( $v = 16$ ) nas três primeiras sondas de linhas de base, selecionando consistentemente faces brancas na presença de palavras positivas e faces negras na presença de palavras negativas. Já um participante (P6) apresentou índice de viés igual a 8, selecionando as faces negras para os adjetivos negativos e o abstrato para o positivo. Após a intervenção, todos os participantes desse grupo deixaram de selecionar sistematicamente faces negras diante de palavras negativas.

Na primeira sonda pós-intervenção (Sonda 4), três participantes (P1, P3 e P4) apresentaram redução imediata do viés, com índice igual a zero. P2 apresentou redução parcial ( $v = 4$ ), enquanto P5 e P6 apresentaram inversão do padrão inicial ( $v = -8$ ), selecionando faces negras na presença de palavras positivas em todas as tentativas. Os padrões observados para P2, P5 e P6 mantiveram-se estáveis nas sondas subsequentes e nos seguimentos de 15, 30 e 45 dias.

Nas sondas posteriores, observou-se alguma variabilidade individual. P1 e P3 mantiveram índice zero na Sonda 5; P2 e P4 apresentaram pequeno aumento do índice. Em sondas subsequentes e nos seguimentos, P3 manteve um padrão estável de baixo viés, enquanto P1 apresentou flutuações com tendência à redução do índice de viés racial ao longo do tempo. P4 apresentou aumento gradual do índice de viés nos seguimentos de 15, 30 e 45 dias. A Figura 1 apresenta os índices de viés racial ao longo das sondas de linha de base, pós-intervenção e seguimento para esses participantes.

Sete participantes (P7–P13) apresentaram pequenas flutuações na linha de base e, por esse motivo, foram expostos a quatro sondas pré-intervenção. P7 apresentou tendência crescente de viés racial ao longo das sondas (12, 14, 16 e 16). Os demais participantes (P8–P13) apresentaram redução do viés na segunda sonda, seguida de estabilidade nas sondas subsequentes.

Todos os participantes desse grupo deixaram de relacionar faces negras a palavras negativas após a intervenção. Nas sondas pós-intervenção e nos seguimentos, observaram-se padrões distintos. P7 e P9 mantiveram índice de viés moderado ( $v = 8$ ) de forma estável nas sondas 5, 6 e 7 e nos seguimentos de 15,

30 e 45 dias. P8 apresentou redução consistente do viés, com índices negativos nas sondas subsequentes ( $v = -2$ ) e valores próximos de zero nos seguimentos. P12 apresentou viés próximo de zero ou negativo na maioria das sondas e seguimentos. Os participantes P10, P11 e P13 exibiram maior variabilidade ao longo do tempo, com oscilações entre índices negativos, moderados e elevados nas sondas pós-intervenção e nos seguimentos, sem um padrão único de manutenção. Finalmente P14 e P15 apresentaram índice de viés racial de 16 em todas as sondas antes e depois da intervenção, bem como nos seguimentos. A Figura 1 apresenta os índices de viés racial ao longo das sondas e dos períodos de seguimento para esses participantes.



### Teste Relacional Preferências com Duas Comparações

Treze participantes apresentaram índice máximo de viés racial ( $b = 16$ ) em todas as sondas realizadas antes da intervenção. Esse padrão indica a seleção sistemática de faces negras na presença de adjetivos negativos e de faces brancas na presença de adjetivos positivos. Apenas dois participantes apresentaram índices ligeiramente inferiores, porém ainda compatíveis com o critério de seleção, sendo superiores a 10 pontos nas sondas pré-intervenção para P8 e superiores a 12 pontos para P11.

Após a intervenção, dez participantes demonstraram mudanças no padrão de seleção. Três deles (P5, P8 e P12) apresentaram índices negativos de viés racial ( $b < 0$ ) nas três sondas subsequentes à intervenção. Esse resultado indica uma inversão do padrão de resposta, caracterizada pela maior seleção de faces negras na presença de adjetivos positivos e de faces brancas na presença de adjetivos negativos. P5 manteve índices negativos estáveis nos seguimentos de 15, 30 e 45 dias. P8 apresentou aumento progressivo do índice, passando de  $-8$  para 8 nos seguimentos de 15, 30 e 45 dias; ainda assim, os valores permaneceram inferiores aos observados antes da intervenção. De forma semelhante, P12 apresentou índice igual a 2 no seguimento de 15 dias e índice 4 nos seguimentos subsequentes, representando aumento em relação às sondas imediatamente pós-intervenção (ver Figura 2).

Os participantes P3 e P6 apresentaram índice de viés igual a zero nas sondas pós-intervenção e mantiveram esse padrão estável nos seguimentos de 15, 30 e 45 dias, indicando ausência de preferência diferencial entre faces brancas e negras.

Alguns participantes apresentaram redução do índice de viés racial, mas mantiveram um padrão ainda favorável às faces brancas. P10, por exemplo, alternou entre índices de 2 e 8 ao longo das sondas pós-intervenção e dos seguimentos. De modo semelhante, P13 apresentou oscilações entre índices 0 e 8 tanto nas sondas pós-intervenção quanto nos seguimentos.

Finalmente, três participantes apresentaram padrões mistos de resposta, alternando entre índices negativos e positivos ao longo do tempo. P1 apresentou índices negativos ( $b < 0$ ) na sonda 7 e nos seguimentos de 15 e 30 dias, bem como índices positivos ( $b > 0$ ) nas sondas 4 e 5, e índice zero no seguimento de 45 dias. Já P11 apresentou índice negativo na sonda 5 e nos seguimentos de 15 e 45 dias, enquanto apresentou índices positivos nas sondas remanescentes e no seguimento de 30 dias.

Cinco participantes (P4, P7, P9, P14 e P15) não demonstraram mudanças no padrão de seleção após a intervenção, mantendo a seleção de faces negras na presença de adjetivos negativos e de faces brancas na presença de adjetivos positivos. No entanto, após a intervenção, todos os participantes supracitados demonstraram maior hesitação ou desconforto durante a tarefa, possivelmente em função da ausência de uma terceira alternativa de resposta, um efeito que não foi observado antes da intervenção.

Cabe destacar que, embora P12 tenha apresentado redução imediata do índice de viés, atingindo valor zero na sonda 4, nas sondas subsequentes e nos seguimentos o índice voltou a níveis elevados, mantendo-se em  $b = 16$ . A Figura 2 apresenta os índices de viés racial ( $b$ ) obtidos no Teste Relacional com Duas Escolhas Forçadas ao longo das sondas de linha de base, das sondas pós-



### **Self-Assessment Manikin**

Treze participantes avaliaram faces negras negativamente (pontos 1–2) e faces brancas positivamente (pontos 4–5) no SAM antes da intervenção. Os participantes P6 e P10 diferiram desse padrão, avaliando as faces negras de forma neutra ( $M = 3$ ) e as faces brancas de forma positiva ( $M = 5$ ).

Após a intervenção, treze participantes apresentaram mudanças nas avaliações afetivas, com exceção de P14 e P15, que mantiveram o padrão observado na linha de base. Os participantes P1 e P11 passaram a avaliar as faces negras de forma neutra ( $M = 3$ ) após a intervenção, mantendo esse padrão tanto nas sondas pós-intervenção quanto nos seguimentos, enquanto as avaliações das faces brancas permaneceram positivas.

Seis participantes (P2, P3, P7, P8, P12 e P13) passaram a avaliar positivamente tanto faces negras quanto brancas nas sondas imediatamente após a intervenção. No entanto, nos seguimentos de 15, 30 e 45 dias, esses participantes passaram a avaliar as faces negras como neutras, enquanto as avaliações das faces brancas permaneceram positivas.

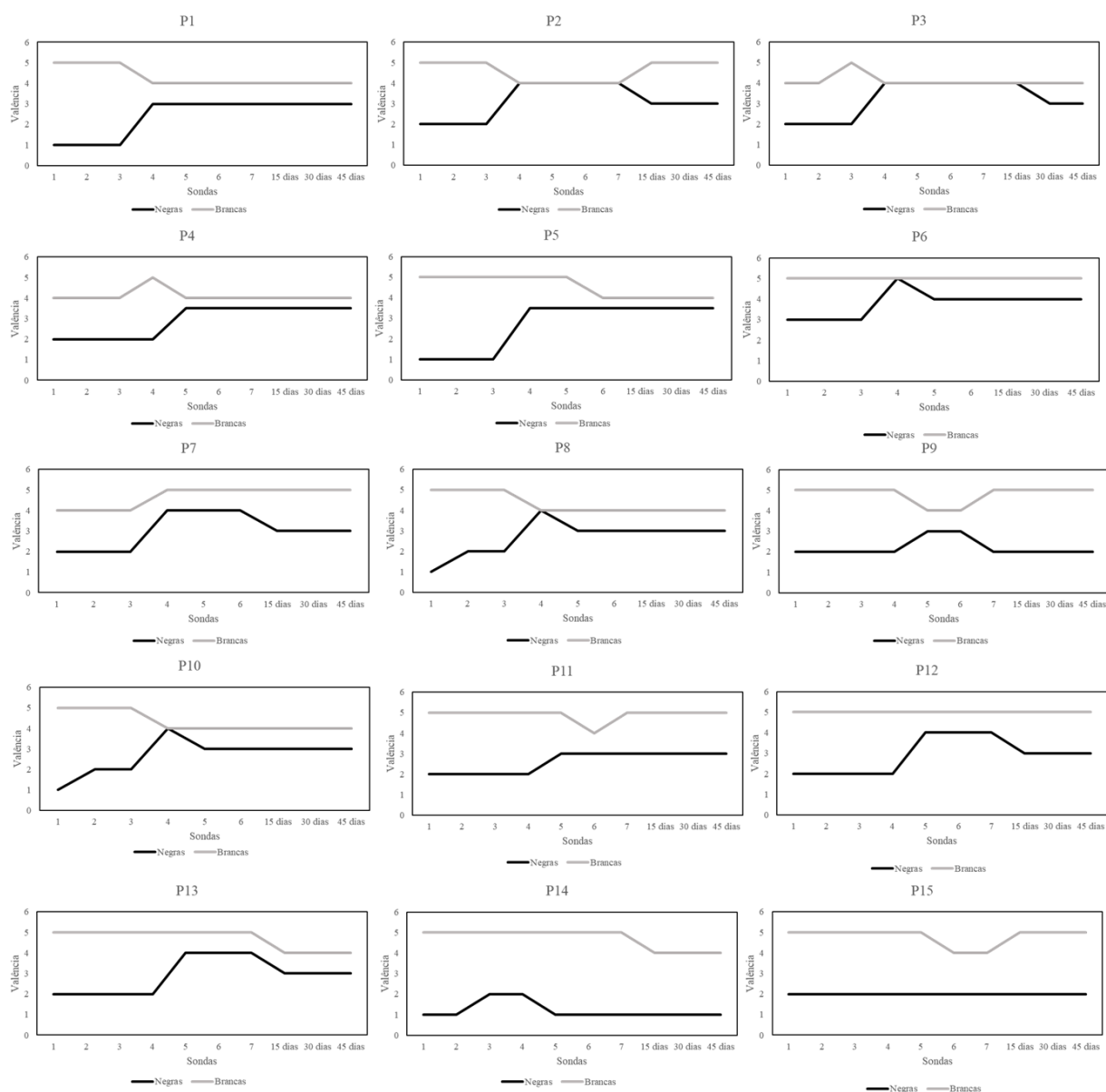
Dois participantes (P6 e P10) apresentaram avaliações predominantemente neutras para faces negras em todas as sondas pós-intervenção e nos seguimentos. Para esses participantes, observou-se diferenciação em função do gênero do estímulo: faces negras femininas foram avaliadas de forma positiva ( $M = 4$ ), enquanto faces negras masculinas foram avaliadas como neutras ( $M = 3$ ). As faces brancas, por sua vez, mantiveram avaliações positivas em todas as sondas.

P6 e P10 também apresentaram, em algumas sondas, avaliações positivas de faces negras ( $M = 4$ ) imediatamente após a intervenção e nos seguimentos, mantendo avaliações positivas para faces brancas ao longo de todo o procedimento.

Finalmente, o participante P9 avaliou as faces negras como neutras em duas sondas subsequentes à intervenção, mas voltou a avaliá-las de forma negativa nas sondas seguintes e nos períodos de seguimento. A Figura 3 apresenta as avaliações afetivas atribuídas às faces negras e brancas pelos participantes na escala SAM, domínio valência.

**Figura 3**

*Valência atribuída às faces negras e brancas no SAM*



### Teste da Boneca Modificado

No Teste da Boneca, as respostas dos participantes foram registradas considerando cinco figuras de bonecas fisicamente idênticas, que diferiam exclusivamente quanto ao tom de pele, organizadas em um contínuo do mais claro

ao mais escuro: bonecas 1 e 2 (brancas), boneca 3 (parda) e bonecas 4 e 5 (negras).

Nas sondas pré-intervenção, diante das perguntas “*Alguma você escolheria para ser sua amiga?*”, todos os participantes selecionaram exclusivamente as bonecas brancas (1 ou 2) e não selecionaram as bonecas negras (4 ou 5). A boneca parda (3) não foi selecionada em nenhuma tentativa. Esse padrão manteve-se consistente ao longo das sondas de linha de base, indicando controle sistemático do tom de pele sobre as escolhas sociais e de preferência, com viés desfavorável em relação às bonecas negras.

Diante da pergunta “*Alguma você escolheria para atividade para nota?*”, todos os participantes selecionaram as bonecas brancas em todas as sondas pré-intervenção. Em relação à pergunta “*Alguma você escolheria para o time da educação física?*”, doze participantes selecionaram bonecas brancas e não selecionaram bonecas negras, enquanto três participantes (P2, P4 e P15) selecionaram bonecas negras e não selecionaram bonecas brancas. Por fim, diante da pergunta “*Alguma você desconfia que roubaria suas coisas?*”, todos os participantes selecionaram bonecas negras (4 ou 5) e não selecionaram bonecas brancas (1 ou 2).

Após a intervenção, nove participantes apresentaram algum grau de mudança nos padrões de seleção, ainda que com variabilidade entre categorias e sondas. Um padrão particularmente consistente foi observado nos participantes P3, P6, P8 e P11, que passaram a selecionar “*todas*” as bonecas como escolha para as categorias “*amiga*”, “*atividade para nota*” e “*time da educação física*”, bem como a responder “*nenhuma*” como não escolha. Adicionalmente, diante da

pergunta “*roubaria suas coisas*”, esses participantes passaram a selecionar “*nenhuma*” como escolha e “*todas*” como não escolha. Esse padrão manteve-se estável nos seguimentos de 15, 30 e 45 dias. A Tabela 1 apresenta a moda das respostas desses participantes, calculada separadamente para as sondas pré- e pós-intervenção.

**Tabela 1.**

*Moda das respostas no Teste da Boneca Adaptado por participante, atividade e tipo de resposta (pré- e pós-intervenção).*

ID	Tipo de Resposta	Amiga		Atividade para nota		Educação Física		Roubaria suas coisas	
		Pré-teste	Pós-teste	Pré-teste	Pós-teste	Pré-teste	Pós-teste	Pré-teste	Pós-teste
P3	Escolhe	2	Todas	1	Todas	1	Todas	1	Todas
	Não Escolhe	4	Nenhuma	4	Nenhuma	4	Nenhuma	4	Nenhuma
P6	Escolhe	2	Todas	1	Todas	2	Todas	5	Todas
	Não Escolhe	5	Nenhuma	5	Nenhuma	5	Nenhuma	1	Nenhuma
P8	Escolhe	2	Todas	2	Todas	2	Todas	2	Todas
	Não Escolhe	5	Nenhuma	5	Nenhuma	5	Nenhuma	5	Nenhuma
P11	Escolhe	1	Todas	1	Todas	1	Todas	1	Todas
	Não Escolhe	4	Nenhuma	4	Nenhuma	4	Nenhuma	4	Nenhuma

*Nota.* Bonecas 1 e 2 são brancas; boneca 3 é parda; bonecas 4 e 5 são pretas. As bonecas diferiam exclusivamente quanto ao tom de pele, organizadas do mais claro ao mais escuro: bonecas 1 e 2 (brancas), boneca 3 (parda) e bonecas 4 e 5 (pretas). “Todas” indica que o participante apontou para todas as bonecas; “Nenhuma” indica ausência de indicação. Os valores apresentados correspondem à moda das respostas emitidas ao longo das sondas pré-intervenção e pós-intervenção, separadamente para as categorias *Escolhe* e *Não escolhe*. A apresentação por moda visa capturar o padrão de resposta mais frequente em uma variável categórica, permitindo identificar mudanças funcionais no controle exercido pelo tom de pele sobre as escolhas sociais.

Quatro participantes (P1, P2, P5 e P12) apresentaram evidências de generalização no Teste da Boneca, porém com maior variabilidade e manutenção parcial do controle racial. Padrões individuais de mudança evidenciam que a generalização observada nesses participantes ocorreu de forma parcial e variável ao longo das sondas pós-intervenção, o que se reflete na análise da moda

apresentada na Tabela 2. No caso de P1, observou-se, ao longo das sondas pós-intervenção e dos períodos de seguimento (15, 30 e 45 dias), alternância entre respostas inclusivas, com indicação de “todas” as bonecas como escolha nas categorias “*amiga*”, “*atividade para nota*” e “*time da educação física*”, e padrões ainda sob controle racial, nos quais a boneca preta de tonalidade mais clara (4) foi indicada como escolha e a boneca preta mais escura (5) como não escolha. Para a pergunta “*roubaria suas coisas*”, manteve-se predominantemente o padrão inverso, com indicação da boneca 5 como escolha e da boneca 4 como não escolha. Apesar da variabilidade entre sondas, a moda pós-intervenção indica ampliação das escolhas para bonecas negras mais claras, com manutenção da exclusão da boneca de tonalidade mais escura.

Em P2, a análise das sondas pós-intervenção mostrou deslocamento consistente do padrão de escolha para a boneca parda (3) e, em algumas categorias, para a boneca preta de tonalidade mais clara (4), especialmente em “*amiga*”, “*atividade para nota*” e “*time da educação física*”. Contudo, a boneca preta mais escura (5) permaneceu predominantemente como não escolhida nessas categorias e como a principal indicada na pergunta “*roubaria suas coisas*”, padrão que se manteve nos seguimentos. A moda pós-intervenção reflete, assim, expansão parcial das escolhas sociais sem eliminação completa do controle aversivo associado ao tom de pele mais escuro.

Por fim, P12 exibiu mudanças específicas por categoria. Nas sondas pós-intervenção, a boneca parda (3) passou a ser a escolha predominante para “*atividade para nota*”, enquanto, para “*time da educação física*”, observou-se deslocamento para a boneca preta de tonalidade mais clara (4), com exclusão da

boneca branca. Quanto à opção “roubaria suas coisas”, o participante seguiu selecionando a boneca preta de tonalidade mais escura. (5) Entretanto, nas demais categorias sociais, o padrão modal permaneceu semelhante ao observado na linha de base, e nos seguimentos houve retorno parcial às escolhas brancas em algumas atividades.

**Tabela 2**

*Moda das escolhas e não escolhas no Teste da Boneca para participantes com generalização parcial*

ID	Tipo de Resposta	Amiga		Atividade para Nota		Educação Física		Roubaria suas coisas	
		Pré-teste	Pós-teste	Pré-teste	Pós-teste	Pré-teste	Pós-teste	Pré-teste	Pós-teste
P1	Escolhe	2	4	2	4	2	4	2	4
	Não Escolhe	5	5	4	5	4	1	4	5
P2	Escolhe	2	2	2	3	2	3	5	5
	Não Escolhe	4	1	4	5	4	5	2	3
P5	Escolhe	2	3	2	4	2	3	5	5
	Não Escolhe	5	5	5	5	5	5	2	2
P12	Escolhe	2	2	2	3	2	4	5	5
	Não Escolhe	4	5	4	4	4	2	2	2

*Nota.* As bonecas diferiam exclusivamente quanto ao tom de pele, organizadas do mais claro ao mais escuro: bonecas 1 e 2 (brancas), boneca 3 (parda) e bonecas 4 e 5 (pretas). Os valores apresentados correspondem à moda das respostas emitidas ao longo das sondas pré-intervenção e pós-intervenção, separadamente para as categorias *Escolhe* e *Não escolhe*. A apresentação por moda visa capturar o padrão de resposta mais frequente em uma variável categórica, permitindo identificar mudanças funcionais no controle exercido pelo tom de pele sobre as escolhas sociais.

Finalmente, sete participantes (P4, P7, P9, P10, P13, P14 e P15) não demonstraram mudanças no padrão de seleção após a intervenção, mantendo a escolha de bonecas brancas para as categorias “*amiga*”, “*atividade para nota*” e “*educação física*”, bem como a não escolha das bonecas negras, e preservando o padrão inverso para a pergunta “*roubaria suas coisas*”.

### **Análise Integrativa dos Efeitos da Intervenção e Convergência entre Medidas**

A análise integrativa dos efeitos da intervenção e da convergência entre medidas revelou três padrões distintos de mudança individual, diferenciados pelo grau de reorganização relacional, pela extensão da generalização para contextos sociais e pela persistência de controles raciais historicamente estabelecidos.

Um primeiro grupo, composto pelos participantes P3, P6, P8 e P12 apresentou alta convergência entre todas as medidas avaliadas. Esses participantes formaram e mantiveram classes de equivalência, exibiram redução consistente do viés nos testes relacionais com duas e três comparações, demonstraram neutralização ou reversão da valência afetiva na escala SAM e generalização funcional no Teste da Boneca. Nesse padrão, observou-se enfraquecimento simultâneo das relações Negro–Negativo e Branco–Positivo, bem como da estrutura relacional de oposição entre escolher e não escolher, indicando elevada flexibilidade relacional.

Um segundo conjunto de participantes, P1, P2, P5 e P12, apresentou mudanças robustas em medidas relacionais e afetivas, porém com generalização parcial para escolhas sociais, conforme sintetizado na Tabela 3. Nesses casos, observou-se redução do viés relacional negativo e neutralização afetiva das faces negras, frequentemente acompanhadas por expansão das escolhas para estímulos pardos e pretos de tonalidade mais clara. Ainda assim, relações Branco–Positivo permaneceram funcionalmente relevantes, sobretudo em tarefas de preferência social, assim como a manutenção de relações de oposição entre escolher e não escolher. De forma consistente, estímulos de tonalidade mais escura continuaram a exercer controle aversivo residual, sugerindo coexistência e

competição entre as novas relações treinadas e repertórios culturais previamente estabelecidos.

Por fim, um terceiro padrão foi observado nos participantes P4, P7, P9, P10, P13, P14 e P15, caracterizando alta resistência à mudança. Nesse grupo, a intervenção produziu efeitos limitados ou restritos às tarefas diretamente relacionadas ao treino, sem generalização consistente para contextos sociais. Nos participantes P14 e P15, a ausência de formação de classes de equivalência foi acompanhada pela manutenção integral do padrão de linha de base.

A Tabela 3 apresenta uma síntese integrativa dos efeitos da intervenção e da convergência entre as diferentes medidas utilizadas, permitindo identificar padrões individuais de mudança em função do grau de reorganização relacional, da generalização para contextos sociais e da persistência de controles raciais historicamente estabelecidos.

### Tabela 3

*Síntese integrativa dos efeitos da intervenção e da convergência entre medidas por participante*

ID	Intervenção	Teste Relacional		Generalização		Síntese do Padrão Individual
	AC e CA (Equivalência)	3 Comparações	2 Comparações	Escala SAM (Valência)	Teste da Boneca	
P1	✓	✓	✓	✓	●	Convergência entre medidas; generalização parcial com controle aversivo residual para tons mais escuros.
P2	✓	✓	×	✓	●	Expansão gradual das classes de escolha para tons pardo e preto claro, com manutenção de relações Branco-Positivo e de

P3	✓	✓	✓	✓	✓	✓	controle aversivo para tons mais escuros. Alta convergência entre medidas; generalização funcional consistente. Neutralização da relação Negro–Negativo, sem
P4	✓	✓	⊙	✓	✓	✗	generalização social, com manutenção de relações de oposição Branco–Escolhe vs. Negro–Não escolhe. Mudança relacional e afetiva robusta; generalização social
P5	✓	✓	✓	✓	✓	⊙	parcial, com manutenção do controle aversivo para tons mais escuros. Alta convergência entre medidas; generalização funcional consistente.
P6	✓	✓	✓	✓	✓	✓	Mudança funcional sob controle das classes de equivalência, com redução do viés relacional negativo, sem transferência para contextos de escolha preferencial.
P7	✓	✓	✗	✓	✓	✗	Alta convergência entre medidas; generalização funcional.
P8	✓	✓	✓	✓	✓	✓	Manutenção do controle Branco–Positivo, com relações de oposição entre escolher e não escolher, e mudanças transitórias não mantidas.
P9	✓	✓	✗	⊙	⊙	✗	Reorganização relacional e neutralização afetiva, com manutenção de um padrão de
P10	✓	✓	✓	✓	✓	✗	

P11	✓	✓	✓	✓	✓	oposição entre escolha e não escolha em contextos de preferência social. Alta convergência entre medidas; generalização funcional consistente.
P12	✓	✓	✓	✓	●	Alta convergência entre medidas, com reorganização relacional e afetiva consistente; generalização parcial para escolhas sociais, com manutenção residual de controle racial em alguns contextos.
P13	✓	✓	●	✓	×	Mudança relacional e afetiva sem generalização para preferências sociais.
P14	×	×	×	×	×	Não formação de classes; manutenção do padrão de linha de base
P15	×	×	×	×	×	Não formação de classes; manutenção do padrão de linha de base

*Nota.* ✓ = mudança consistente e/ou mantida nas sondas pós-intervenção e seguimentos; ● = mudança parcial, instável ou dependente do contexto; × = ausência de mudança.

### Validade Social

A validade social da intervenção foi avaliada por meio da percepção de professores, coordenação e direção escolar (N = 4). Os resultados indicaram alta relevância social do estudo, com média de avaliação máxima (M = 5), sugerindo consenso de que o tema do preconceito racial é pertinente e deve ser investigado e discutido no contexto escolar.

Quanto à eficácia e efetividade percebida da intervenção, os professores atribuíram escores elevado para mudanças de comportamentos verbais negativos

dos alunos ( $M = 5$ ); um escore alto para mudanças na forma de interação dos alunos com adultos e crianças negras ( $M=4$ ). Esses dados indicam que as mudanças avaliativas foram observadas no cotidiano da sala de aula.

Nos comentários qualitativos, foi relatado que o participante P5 apresentava previamente comportamentos verbais raciais ofensivos direcionados a um colega; após a intervenção, tais comportamentos deixaram de ocorrer, mantendo-se ausentes até o momento de preenchimento do questionário. Esse relato foi corroborado pelo coordenador pedagógico, que também mencionou a ausência de novas intercorrências envolvendo o participante.

O coordenador e o diretor atribuíram escores médios-alto à efetividade da intervenção ( $M=3.5$ ), atribuindo escore 4 às mudanças de comportamentos verbais negativos; 3 para mudanças no comportamento de interação com pares e adultos negros; e 4 à redução do viés racial. Contudo, esses respondentes ponderaram que não acompanham rotineiramente as interações em sala de aula, o que limita a observação direta de mudanças comportamentais. Essa diferença entre informantes é consistente com o grau de exposição aos comportamentos-alvo e não indica ausência de efeito.

### **Discussão**

O presente estudo avaliou a generalidade dos efeitos de um procedimento baseado na formação de classes de equivalência envolvendo palavras estereotípicas de valência positiva e negativa, examinando em que medida as mudanças produzidas se estenderam a diferentes respostas e medidas relacionadas ao viés racial em crianças. Os resultados, de modo geral, replicam achados prévios quanto à eficácia do procedimento (Mizael et al., 2016, 2020;

Passarelli, Roche, et al., 2024), e avançam ao demonstrar que seus efeitos podem se generalizar, ainda que de forma parcial e variável, para medidas avaliativas distintas e contextos diferentes daqueles diretamente treinados. Essas alterações, no entanto, ocorreram de forma heterogênea entre participantes e entre tarefas, um padrão consistente com a concepção do preconceito racial como um fenômeno comportamental complexo, composto por múltiplas classes de respostas sensíveis a diferentes contextos e contingências (Passarelli et al., 2023; ver também Passarelli, Roche et al., 2024; Passarelli, Roche, et al., 2025).

A elevada taxa de formação de classes de equivalência observada neste estudo, bem como a manutenção dessas classes ao longo dos períodos de seguimento, replica e estende achados prévios dessa linha de pesquisa (Mizael et al., 2016, 2020; Passarelli, Roche, et al., 2024). Esse resultado é particularmente relevante à luz da literatura que aponta que estímulos com funções pré-experimentais afetivamente carregadas tendem a dificultar a formação de classes de equivalência, sobretudo quando envolvem relações inconsistentes com histórias relacionais previamente estabelecidas (Haydu et al., 2015; Moxon et al., 1993; Watt et al., 1991). O sucesso observado no presente estudo sugere, portanto, que os parâmetros adotados, incluindo a estrutura do treino, a organização simples-para-complexo e treino misto (Adams et al., 1993; Mizael et al., 2020) podem ter atuado como atenuadores dessas dificuldades. Ademais, a faixa etária das crianças constitui um fator relevante a ser considerado, uma vez que histórias relacionais associadas a palavras estereotípicas, embora já presentes, são presumivelmente menos densas e menos rigidamente consolidadas na infância, o que pode aumentar a sensibilidade a reorganizações

funcionais produzidas por procedimentos baseados em equivalência (Levy & Hughes, 2009; Lai et al., 2014, 2016).

Embora a formação de classes tenha sido consistente para a maioria dos participantes, dois deles apresentaram um padrão distinto, que pode ser interpretado à luz do conceito de resistência à mudança comportamental, conforme descrito pelo modelo de *momentum comportamental* (Nevin, 1992; Nevin & Grace, 2000; ver também Belisle & Dixon, (2020) para a aplicação desta noção ao responder relacional). Esses participantes apresentaram desempenho adequado nos treinos das relações condicionais básicas, indicando aprendizagem eficiente, mas não demonstraram emergência consistente das relações derivadas que implicavam a reorganização de funções avaliativas previamente estabelecidas. Evidências prévias indicam que relações simbólicas com história pré-experimental densa ou carga afetiva podem interferir na formação e reorganização de classes de equivalência, mesmo sob parâmetros de treino adequados (Haydu et al., 2015; Moxon et al., 1993; Watt et al., 1991). De forma consistente, a ausência de mudanças de viés racial ocorreu paralelamente à não formação das classes, reforçando a relação funcional entre a emergência de classes de equivalência e os efeitos de generalização observados.

As mudanças observadas após a intervenção foram mais consistentes nos testes relacionais computadorizados, os quais avaliam respostas sob controle direto das relações ensinadas durante o treino de equivalência. Esse padrão é compatível com análises conceituais que enfatizam que tarefas com contingências mais restritas e explicitamente programadas tendem a captar de forma mais direta o controle relacional estabelecido experimentalmente, em comparação a medidas

que envolvem maior variabilidade contextual e múltiplas fontes de controle (Barnes-Holmes et al., 2010; Critchfield, 2011; Mace & Critchfield, 2010).

A comparação entre os dois testes relacionais, idênticos em estrutura e distintos apenas pelo número de alternativas de resposta, revelou que a intervenção foi mais eficaz em enfraquecer a relação negro–negativo do que em reorganizar a relação branco–positivo. No teste com três comparações, a inclusão de um estímulo abstrato, menos carregado por história pré-experimental, pode ter favorecido respostas sob controle das relações recentemente treinadas, resultando em mudanças mais consistentes. Esse padrão é compatível com efeitos observados em estudos anteriores, nos quais procedimentos baseados em equivalência também produziram reorganizações avaliativas mais claras em contextos de escolha relacional (Mizael et al., 2016, 2020). Em contraste, no teste com duas comparações, em que apenas alternativas raciais estavam disponíveis, a preferência por faces brancas permaneceu saliente para parte dos participantes, especialmente em contextos de ambiguidade decisória. Esse padrão é consistente com a literatura sobre viés racial, que indica maior resistência à mudança de repertórios pró-branco em comparação àqueles explicitamente negativos dirigidos a grupos racializados (Correll et al., 2002; Bigler & Liben, 2007; Nesdale et al., 2005; Lai et al., 2014, 2016).

Do ponto de vista analítico-comportamental, esses achados são compatíveis com a noção de que diferentes relações simbólicas apresentam distintos graus de força histórica e resistência à mudança, sendo selecionadas diferencialmente em função das propriedades do contexto de teste e das alternativas disponíveis (Belisle & Dixon, 2020; Nevin, 1992; Nevin & Grace, 2000;

Hughes & Barnes-Holmes, 2011), reforçando a necessidade de problematizar preferências raciais mesmo na ausência de avaliações negativas explícitas (Passarelli et al., 2023; Passarelli, Roche et al., 2024; Roche et al., 2025).

Os resultados obtidos na avaliação de valência replicam achados prévios dessa linha de pesquisa, indicando que procedimentos baseados na formação de classes de equivalência podem alterar avaliações afetivas relacionadas a estímulos raciais (Mizael et al., 2016, 2020; Passarelli, Roche et al., 2024; Passarelli et al., 2026, em preparação). Para a maioria dos participantes, a intervenção foi seguida por uma mudança consistente nas avaliações das faces negras, que deixaram de ser avaliadas negativamente e passaram a ser avaliadas como neutras ao longo das sondas pós-intervenção e dos períodos de seguimento. Esse padrão é compatível com um processo de transferência de valência, observado em procedimentos de condicionamento avaliativo (ver Lai et al., 2014, 2016), no qual a função avaliativa negativa previamente associada às faces negras foi atenuada.

É importante destacar que essa neutralização afetiva se mostrou suficiente para mitigar o viés antinegro; contudo, as avaliações das faces brancas permaneceram predominantemente positivas e superiores às atribuídas às faces negras, indicando a manutenção de um controle diferencial baseado na raça. Esse padrão converge com os resultados do Teste Relacional de Preferência com Duas Escolhas, no qual, em contextos de escolha forçada, estímulos com maior valência positiva tendem a ser preferidos (Santos & de Rose, 2018, 2019), e com os resultados do Teste da Boneca, no qual, mesmo quando bonecas negras foram escolhidas em avaliações positivas, as de tonalidade de pele mais clara

receberam resultados mais favoráveis, enquanto a boneca de tonalidade mais escura permaneceu, em geral, em posição de exclusão ou associada a avaliações negativas, sugerindo que os efeitos da intervenção se distribuíram de forma diferencial entre as tonalidades de pele.

Assim, os dados da Escala SAM sugerem que a intervenção pode ter reduzido o controle aversivo previamente exercido por estímulos raciais negros. Contudo, esses resultados não excluem a possibilidade de que uma valência positiva diferencial atribuída ao endogrupo branco também tenha contribuído para o padrão observado, especialmente considerando que relações historicamente reforçadas de positividade associadas à branquitude podem permanecer salientes. Evidências provenientes de estudos com tarefas relacionais implícitas indicam que avaliações positivas do endogrupo podem ocorrer independentemente da presença de avaliações negativas explícitas do exogrupo, variando conforme a população e o contexto de mensuração. Nesse sentido, os efeitos observados podem refletir uma combinação de redução do controle aversivo associado ao exogrupo e manutenção parcial de relações apetitivas associadas ao endogrupo, sobretudo em tarefas menos estruturadas e que exigem comparações diretas entre estímulos sociais (Barnes-Holmes, Murphy et al., 2010; Drake et al., 2010).

Os participantes que apresentaram maior flexibilidade no Teste da Boneca (P3, P6, P10 e P11) também foram aqueles que exibiram padrões mais consistentes de mudança ao longo das diferentes medidas avaliativas. Todos formaram classes de equivalência com manutenção até 45 dias, deixando de relacionar sistematicamente faces negras a estímulos negativos nos testes relacionais, o que indica reorganização efetiva do controle simbólico exercido pela

raça (Mizael et al., 2016, 2020; Passarelli, Roche, et al., 2024). Ademais, esses participantes apresentaram neutralização ou valência positiva nas avaliações afetivas das faces negras na Escala SAM, um padrão compatível com transferência de função afetiva previamente descrita na literatura (Dougher et al., 1994; de Rose et al., 1988; Santos & de Rose, 2018, 2019). De forma convergente, no Teste da Boneca modificado, esses participantes eliminaram padrões de exclusão social e estereotipia racial, passando a incluir todas as bonecas nas atividades propostas e a rejeitar atribuições associadas à criminalidade. Esse conjunto de resultados sugere que, para esse subgrupo, a intervenção produziu mudanças amplas e coerentes entre medidas relacionais, afetivas e sociais, caracterizando um padrão robusto de flexibilidade relacional e generalização funcional para contextos sociais menos estruturados, o que não foi observado em estudos prévios dessa linha de pesquisa (Mizael et al., 2016, 2020; Passarelli, Roche, 2024; Passarelli, 2026).

Em quatro participantes, observou-se evidência de generalização parcial, caracterizada por aumento da variabilidade nas respostas ao longo das sondas, mas com alternância entre padrões mais inclusivos e a manutenção de controle racial previamente estabelecido. Esses participantes, embora tenham deixado de relacionar sistematicamente faces negras a estímulos negativos nos testes relacionais e apresentado neutralização afetiva nas avaliações de valência, continuaram a excluir de forma recorrente a boneca de tonalidade mais escura nas escolhas por atividades sociais e a indicá-la como a mais provável de “roubar suas coisas” no Teste da Boneca. Esse padrão sugere que a reorganização relacional promovida pela intervenção foi suficiente para atenuar relações aversivas

explícitas, mas não para eliminar completamente funções discriminativas negativas associadas ao tom de pele mais escuro, evidenciando um efeito compatível com o fenômeno do colorismo (Maddox, 2004; Hunter, 2007). Assim, novas relações treinadas passaram a coexistir com repertórios historicamente estabelecidos, sendo evocadas diferencialmente conforme as propriedades da tarefa e o grau de estruturação do contexto avaliativo.

Em contraste, seis participantes não apresentaram mudanças no Teste da Boneca, mantendo padrões estáveis de exclusão racial ao longo de todas as sondas, incluindo dois que não formaram classes de equivalência. Para esses participantes, embora alguns tenham apresentado reduções pontuais no viés nos testes relacionais mais estruturados ou neutralização parcial nas avaliações afetivas, tais mudanças não se estenderam às escolhas por atividades sociais, nas quais o controle racial permaneceu consistente. Esse padrão sugere que as relações treinadas foram insuficientes para reorganizar repertórios sociais mais amplos e historicamente consolidados. À luz do modelo de momentum comportamental, esse conjunto de resultados é compatível com maior resistência à mudança, uma vez que repertórios mantidos por histórias densas e socialmente reforçadas tendem a persistir mesmo diante de novas contingências experimentais (Nevin, 1992; Nevin & Grace, 2000). Essa interpretação é consistente com extensões recentes do conceito de momentum ao responder relacional, nas quais a densidade de redes relacionais tem sido associada à resistência à reorganização funcional (Belisle & Dixon, 2020). Assim, a dissociação entre efeitos observados em tarefas relacionais e afetivas e a ausência de mudanças em escolhas sociais explícitas indica que, para esses participantes, seriam

necessários procedimentos adicionais ou contextos de intervenção mais próximos das contingências naturais para promover generalização funcional em situações sociais menos estruturadas.

Adicionalmente, os dados de identificação racial indicam que, embora a maioria das crianças tenha apresentado correspondência entre autoidentificação, fenótipo e classificação parental, uma parcela dos participantes se autoidentificou como branca apesar de apresentar fenótipo pardo e identificação parental convergente. Esse padrão sugere que a identificação racial não é controlada exclusivamente por propriedades físicas, mas constitui um comportamento verbal sensível a contingências sociais, normas culturais e processos de deseabilidade social. Resultados clássicos do Teste da Boneca já indicavam que crianças tendem a atribuir maior valor social e afetivo à brancura, inclusive quando essa categorização não corresponde estritamente ao fenótipo (Clark & Clark, 1947), um achado amplamente replicado e reinterpretado à luz de fenômenos como o colorismo. No contexto brasileiro, marcado por forte miscigenação e fronteiras raciais ambíguas, indivíduos de tonalidade intermediária frequentemente se aproximam simbolicamente da categoria “branco”, em um processo de embranquecimento socialmente reforçado (Nogueira, 2006; Telles, 2004; Munanga, 2004). Estudos sobre colorismo demonstram que pequenas variações no tom de pele estão associadas a diferenças sistemáticas em status social, avaliações interpessoais e autoatribuição racial, com tonalidades mais claras funcionando como estímulos discriminativos associados a maior valor social (Maddox, 2004; Hunter, 2007). Evidências indicam ainda que, desde a infância, a identificação racial é moldada por práticas familiares, escolares e midiáticas, nas

quais a brancura tende a operar como categoria normativa e desejável, influenciando padrões de autoidentificação mesmo na ausência de correspondência fenotípica estrita (Harris, 2006; Camilo et al., 2020). Esses achados ajudam a contextualizar tanto os dados de identificação racial quanto os padrões observados nas tarefas de preferência social, sugerindo que o viés racial se organiza em gradientes sutis de cor da pele, e não apenas em categorias dicotômicas (Ver Mizael, Castro & Dittrich, 2021).

As informações provenientes da avaliação de validade social indicam que a intervenção foi considerada altamente relevante e socialmente significativa pela equipe escolar, reforçando sua pertinência aplicada no contexto educacional. As avaliações positivas atribuídas pelos professores, especialmente no que se refere à redução de comportamentos verbais raciais ofensivos e à melhora nas interações com pares e adultos negros, sugerem que os efeitos observados nas medidas experimentais se estenderam para o cotidiano da sala de aula. Relatos qualitativos, como a ausência de reincidência de comentários raciais ofensivos por parte de um participante específico, são consistentes com a literatura que enfatiza a importância de evidências naturalísticas na avaliação de intervenções voltadas a repertórios sociais complexos (Wolf, 1978; Baer, Wolf, & Risley, 1968).

O acompanhamento sistemático dos participantes ao longo do tempo, aliado ao engajamento ativo da escola na discussão do tema, pode ter contribuído para a manutenção dos efeitos observados, ao favorecer o contato contínuo com contingências sociais naturais coerentes com os objetivos da intervenção. A literatura aponta que esse tipo de suporte contextual é um fator crítico para a generalização e a sustentação de mudanças comportamentais em ambientes

educacionais (Stokes & Baer, 1977; Cooper et al., 2020). As diferenças entre as avaliações dos professores e da equipe gestora são compatíveis com o grau diferencial de exposição aos comportamentos-alvo e não indicam ausência de efeito, mas reforçam a relevância do uso de múltiplos informantes na análise da validade social.

-----

### Considerações Finais do Capítulo

Os resultados deste estudo somam-se às evidências empíricas acumuladas em pesquisas anteriores, indicando que a formação de classes de equivalência envolvendo estímulos socialmente carregados (e.g., faces negras) e estímulos positivos é suficiente para reorganizar, de modo consistente, relações avaliativas negativas dirigidas a faces negras no teste relacional. Nos demais instrumentos, os efeitos foram parciais e variáveis entre participantes, com alguns resultados convergentes e outros sem alteração detectável, o que reforça a necessidade de cautela na generalização das conclusões para além das medidas diretamente sensíveis ao treino de equivalência. Esse achado é particularmente relevante porque o procedimento utilizou palavras estereotípicas, presumivelmente dotadas de história verbal pré-experimental extensa, o que aumenta a probabilidade de conflito relacional durante o treino, mas também aproxima o procedimento das contingências culturais nas quais o viés racial é estabelecido e mantido.

No entanto, assim como observado em estudos anteriores, a reorganização relacional não se estendeu de maneira uniforme a todas as medidas avaliativas. Embora a generalização tenha ocorrido para uma parcela substancial dos participantes, esse efeito não foi observado de forma sistemática em toda a amostra, evidenciando variabilidade interindividual nos efeitos da intervenção.

Esse padrão convida à reflexão sobre as fontes dessa variabilidade. A ausência de generalização plena não aponta necessariamente para limitações do procedimento em si, mas para a interação entre o arranjo experimental e as histórias pré-experimentais dos participantes. Os dados sugerem que o treino baseado em equivalência atua de forma diferencial conforme o repertório relacional prévio: enquanto alguns participantes

demonstram maior flexibilidade e sensibilidade à reorganização simbólica, outros apresentam histórias mais densas e resistentes à mudança.

Sob essa perspectiva, a variabilidade observada não deve ser interpretada como falha, mas como evidência empírica das condições sob as quais a reorganização relacional é mais ou menos provável. Tal como ocorre em outras tecnologias comportamentais, e, de modo análogo, em intervenções farmacológicas, não se espera uniformidade absoluta de efeitos. O fato de o procedimento produzir redução consistente de avaliações negativas sempre que houve formação de classes, e efeitos adicionais em outros domínios para parte dos participantes, reforça sua robustez enquanto ferramenta experimental para o estudo da transformação de funções simbólicas, ainda que seus efeitos dependam criticamente da história e do contexto do indivíduo.

Um padrão particularmente relevante observado neste estudo foi o de participantes que passaram a selecionar todas as bonecas para atividades socialmente positivas e nenhuma para categorias de não escolha ou atribuição negativa. Esse padrão representa mais do que uma simples redistribuição de preferências: indica uma reorganização funcional do controle exercido pela raça no contexto decisório. Contudo, essa reorganização não foi uniforme entre as tonalidades de pele: em alguns participantes, a boneca de tonalidade mais escura continuou associada a avaliações negativas mesmo após a intervenção, sugerindo que o tom de pele operou como estímulo discriminativo de forma gradual e não binária. Nesses casos, o responder inclusivo e não diferencial foi observado apenas parcialmente, circunscrito às bonecas de tonalidades intermediárias, o que reforça a relevância do colorismo como variável a ser sistematicamente investigada em estudos futuros.

Diferentemente de padrões nos quais a redução do viés se manifesta apenas como deslocamento gradual de escolhas, por exemplo, da boneca branca para a parda ou para a negra de tonalidade mais clara, a seleção de todas as bonecas sugere o enfraquecimento da estrutura relacional dicotômica que organizava o repertório em termos de escolher versus não escolher com base na raça. Do ponto de vista relacional, trata-se de um aumento de flexibilidade, no qual categorias sociais previamente organizadas por oposição racial deixam de restringir o repertório de escolha.

Outros participantes apresentaram formas menos robustas, porém ainda relevantes, de flexibilização relacional, deslocando suas escolhas para bonecas de outras raças, ainda que mantendo algum grau de controle racial. Esses padrões intermediários sugerem coexistência entre novas relações emergentes e repertórios historicamente estabelecidos. Mesmo entre participantes que não demonstraram mudanças no Teste da Boneca, mas apresentaram alterações em outras medidas, observa-se indício de flexibilização relacional dependente do contexto de resposta.

Um dado que se repete ao longo desta tese, e que se mantém neste estudo, é que na ausência de formação de classes de equivalência não se observam mudanças funcionais relevantes nas demais medidas. Participantes que não apresentaram emergência relacional mantiveram padrões semelhantes aos da linha de base, indicando resistência à mudança mesmo após exposição ao procedimento. Esses resultados sugerem que algum grau de reorganização relacional é condição necessária para a ocorrência de mudanças avaliativas ou sociais, ainda que não seja suficiente para garantir generalização ampla. A equivalência, nesse sentido, não assegura a mudança, mas sua ausência delimita claramente seus limites. O que se observa é uma relação funcional entre o procedimento e a possibilidade de

mudança, modulada pela resistência comportamental e pela história prévia de cada participante.

Os dados de validade social acrescentam uma dimensão crítica à interpretação dos resultados, ao situarem os efeitos do procedimento em um contexto institucional real e ao longo do tempo. A presença continuada da pesquisadora na escola, o acompanhamento longitudinal dos participantes e a articulação sistemática com a equipe pedagógica ampliaram não apenas a aceitabilidade do procedimento, mas também sua visibilidade funcional no cotidiano escolar. Diferentemente de intervenções pontuais, o contato prolongado permitiu que mudanças comportamentais fossem observadas em interações naturais, sob controle de contingências sociais não programadas experimentalmente.

A atuação conjunta com professores, coordenação e direção favoreceu a integração do procedimento à rotina da escola, reduzindo sua caracterização como uma atividade externa ou artificial. Essa inserção progressiva parece ter aumentado a sensibilidade da equipe para identificar mudanças sutis, porém socialmente relevantes, como a redução de comportamentos verbais racialmente ofensivos e alterações na forma de interação entre as crianças. Sob essa perspectiva, a validade social não decorre apenas dos efeitos produzidos pelo treino, mas da relação estabelecida entre pesquisa, contexto e agentes naturais de reforçamento.

Esses dados reforçam a noção de que a validade social não é uma propriedade intrínseca do procedimento, mas um produto da história de interação entre pesquisadores e comunidade. Intervenções que se estendem no tempo e que envolvem a escola como parceira ativa tendem a produzir efeitos mais observáveis, mais sustentáveis e mais alinhados às práticas educativas, mesmo quando os efeitos experimentais são heterogêneos entre participantes. Assim, o presente estudo sugere que a aproximação contínua com o

contexto escolar constitui um componente funcional relevante para a generalização e a manutenção de mudanças relacionadas a repertórios socialmente sensíveis, como o preconceito racial.

\*\*\*\*\*

## Capítulo 5. Intervenção para Preconceito Racial em Adultos

### Background

A opção inicial por conduzir estudos com crianças fundamentou-se na suposição de que, em comparação a adultos, elas apresentariam uma história pré-experimental menos consolidada, o que tenderia a favorecer maior sensibilidade à mudança relacional.

Adicionalmente, crianças costumam apresentar menor filtragem social do responder, em função do estágio de desenvolvimento e do contato ainda limitado com contingências sociais punitivas associadas à expressão explícita de atitudes socialmente indesejáveis.

Contudo, os dados empíricos acumulados ao longo desta tese indicam que essa expectativa deve ser qualificada. Estudos com crianças entre oito e dez anos já evidenciam padrões consistentes de desejabilidade social e resistência à mudança, aproximando-se, em alguns aspectos, do responder observado em adultos. Mesmo em crianças mais jovens, em torno de seis anos, faixa etária na qual se observou maior correspondência entre medidas e menor influência da desejabilidade social, foram identificados padrões relevantes de resistência à reorganização relacional.

Esses resultados sugerem que, embora a densidade da história pré-experimental seja um fator relevante, ela não atua de forma linear nem garante, por si só, maior flexibilidade comportamental. A resistência à mudança parece emergir precocemente, possivelmente como produto de práticas culturais amplas que organizam o comportamento racial desde os primeiros anos do desenvolvimento. É a partir dessa constatação que se justifica a transição para a investigação de procedimentos voltados explicitamente à promoção de flexibilidade relacional. Os dados indicam que a alteração de valências ou a formação pontual de novas relações não é suficiente para produzir reorganizações funcionais amplas quando o comportamento está sustentado por redes simbólicas densas.

Nesse contexto, destaca-se o *Function Acquisition Speed Test* (FAST), instrumento desenvolvido no âmbito da Teoria das Molduras Relacionais (RFT) utilizado para investigar a fluência com que indivíduos adquirem e emitem respostas relacionais consistentes com estereótipos culturalmente predominantes, em comparação com situações nas quais essas relações são inconsistentes. Diferentemente de medidas que capturam apenas o conteúdo das avaliações, o FAST permite examinar a dinâmica do responder relacional, isto é, a velocidade de aquisição de funções sob diferentes arranjos relacionais.

Assim, o FAST pode ser compreendido não apenas como uma medida da organização das relações verbais de um indivíduo, mas também como um indicador funcional de resistência à mudança ou de flexibilidade relacional. Aquisição rápida em relações consistentes e lentificação acentuada em relações inconsistentes sugerem maior rigidez relacional, ao passo que desempenhos mais simétricos entre condições indicam maior sensibilidade às contingências atuais. O instrumento permite, portanto, tornar observável a competição entre repertórios historicamente estabelecidos e novas exigências contextuais.

Se o FAST tem sido empregado como ferramenta descritiva, ele também suscita uma questão empírica central: o próprio procedimento poderia funcionar como um treino capaz de promover maior flexibilidade relacional? Essa pergunta torna-se particularmente relevante quando considerada em populações adultas, cujas histórias relacionais tendem a ser mais extensas, densas e rigidamente estabelecidas, configurando um cenário mais exigente para testar os limites da reorganização simbólica.

Partindo dessa questão, o estudo que segue foi conduzido no contexto de um Estágio de Pesquisa no Exterior, financiado pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), realizado na Universidade de Maynooth, na Irlanda. A

amostra foi composta por adultos, selecionados por conveniência em função das condições contextuais, temporais e logísticas envolvidas na realização de uma pesquisa desenvolvida no exterior. Cabe destacar que o estágio forneceu as condições institucionais e operacionais necessárias para a execução do experimento, sem implicar alterações nos pressupostos teóricos ou no programa de pesquisa desenvolvido ao longo desta tese. Ademais, considerando que padrões de resistência à mudança também foram observados em estudos com crianças, os resultados obtidos neste contexto podem informar investigações futuras e ser, cautelosamente, estendidos a esse grupo.

O conjunto de medidas foi planejado para captar diferentes dimensões do viés racial. Utilizou-se uma medida explícita indireta, a Escala de Racismo Moderno adaptada para a população irlandesa, voltada à identificação de formas sutis e socialmente aceitáveis de racismo contemporâneo. Como complemento, foi empregada uma medida de valência afetiva, sensível a mudanças avaliativas diretas atribuídas a estímulos raciais.

Por fim, foi utilizada uma tarefa implícita de tomada de decisão sob pressão temporal, o Jogo do Atirador (*Shooter Game*), inspirado no paradigma do Dilema do Policial (*Police Officer's Dilemma Task*). Essa medida implícita exige decisões rápidas de disparo ou não disparo diante de alvos armados e desarmados de diferentes grupos raciais, reduzindo o controle deliberativo e favorecendo o controle por relações previamente estabelecidas. Ao mesmo tempo, oferece um panorama mais próximo de contextos naturalísticos de tomada de decisão, ampliando a validade ecológica da avaliação.

O experimento descrito a seguir emprega diferentes configurações contextuais do FAST como procedimento de treino, com o objetivo de examinar efeitos de prática, flexibilidade relacional e resistência à mudança em múltiplas medidas. Um grupo foi exposto à alternância entre blocos consistentes e inconsistentes, visando promover

flexibilidade relacional; um segundo grupo foi exposto apenas a blocos inconsistentes, reforçando relações contra estereotípicas; um terceiro grupo foi exposto exclusivamente a blocos consistentes, permitindo observar os efeitos da prática reiterada de relações estereotípicas; e um grupo controle foi incluído para isolar os efeitos específicos do treino.

Esse delineamento permite examinar como manipulações contextuais no treino relacional produzem efeitos diferenciais sobre o viés racial, avaliados por meio de medidas explícitas, avaliativas e implícitas. Ao comparar padrões de flexibilidade, resistência à mudança e generalização entre grupos, o estudo busca avançar na compreensão dos processos comportamentais envolvidos na modificação, ou manutenção, do viés racial antinegro.

\*\*\*\*\*

## **Estudo 7 - Mitigando o viés racial em adultos brancos por meio de treino relacional profilático com o Function Acquisition Speed Test (FAST)<sup>12</sup>**

Vieses raciais podem ser avaliados por métodos de autorrelato e por tarefas comportamentais sob pressão temporal, como jogos de tiro simulados baseados na Tarefa do Dilema do Policial (*Police Officer's Dilemma Task* – PODT). Este estudo adota uma abordagem analítico-comportamental para investigar a redução do viés racial nessas medidas, com foco na modificação do controle relacional que orienta decisões rápidas de disparo. A intervenção profilática consistiu na exposição de 116 adultos brancos (idade média = 23,13 anos; 67 mulheres) a diferentes configurações do *Function Acquisition Speed Test* (FAST), organizadas em quatro condições, antes da aplicação de duas medidas de autorrelato e da PODT. As condições visaram estabelecer classes de resposta funcional consistentes com estereótipos raciais para um grupo, inconsistentes com esses estereótipos para outro grupo ou consistentes e inconsistentes alternadas (flexibilidade relacional) para um terceiro grupo, além de uma condição controle sem intervenção. Os participantes da condição consistente com estereótipos apresentaram maior acurácia nas decisões de disparo contra alvos negros armados do que contra alvos brancos armados, bem como maior frequência de falsos alarmes para alvos negros desarmados, padrão também observado no grupo controle. Esses vieses não foram observados nas condições de flexibilidade relacional ou inconsistentes com estereótipos. Os resultados indicam que o treino relacional verbal pode reduzir vieses raciais em decisões de disparo sob pressão temporal, no contexto experimental investigado.

---

<sup>12</sup> <sup>12</sup> Passarelli, D. A., Mizael, T. M., & De Rose, J. C. C. (2026). *Mitigating Racial Bias in a White Sample Using a Prophylactic Functional Response Class Training Method*. Manuscrito no Prelo do Periódico *Journal of Contextual Behavior Science*.

## Introdução

O viés racial influencia o comportamento em contextos socialmente relevantes, incluindo policiamento, saúde, educação e emprego (Greenwald & Krieger, 2006; Nosek et al., 2007). Embora expressões explícitas de preconceito tenham diminuído em muitas sociedades, padrões de resposta racialmente enviesados ainda ocorrem, particularmente sob pressão de tempo ou em condições de ambiguidade (Correll et al., 2002, 2007; Mekawi & Bresin, 2015; Passarelli, Mizael et al., 2024). A partir de uma perspectiva funcional, tais padrões de resposta podem ser compreendidos não como processos inconscientes, mas como respostas relacionais historicamente reforçadas a estímulos socialmente salientes (Barnes-Holmes et al., 2010; Hughes & Barnes-Holmes, 2017; Passarelli, Mizael et al., 2024; Watters et al., 2023). De acordo com essa visão, o viés racial emerge quando classes de estímulos como “negro” se tornam funcionalmente relacionadas a termos avaliativos negativos por meio de extensas histórias de aprendizagem (Mizael et al., 2016; Mizael et al., 2021; Mizael & de Rose, 2017; Passarelli et al., 2023; Passarelli, Roche et al., 2024; Passarelli, Mizael et al., 2024). Essas relações verbais frequentemente persistem entre contextos ou mesmo diante de contingências contraditórias (Gopal et al., 2021; Maina et al., 2018; Tajeu et al., 2018). Assim, a redução desse viés pode depender do aumento da flexibilidade do responder relacional, em vez de simplesmente suprimir ou retrainar relações individuais (Barnes-Holmes et al., 2020; Passarelli, Mizael et al., 2024).

O Teste de Velocidade de Aquisição de Funções (*Function Acquisition Speed Test*, FAST - O'Reilly et al., 2012) é um procedimento analítico-

comportamental que indexa a fluência relativa com que indivíduos adquirem e revertem relações arbitrárias entre estímulos sob contingências de reforçamento (Cummins & Roche, 2020; Watters et al., 2023). Diferentemente de medidas tradicionais baseadas em latência, como o *Implicit Association Test* (IAT; Greenwald et al., 1998), o FAST fornece uma medida direta de “momentum comportamental”, um construto fundamentado no contexto comportamental, que reflete a resistência à mudança em redes verbais (Nevin et al., 1983; Watters et al., 2023). Quando aplicado ao viés racial, o FAST tem sido utilizado para comparar a velocidade de aprendizagem de relações estímulo–estímulo consistentes com estereótipos versus inconsistentes com estereótipos, oferecendo uma forma de quantificar a força de repertórios verbais socialmente reforçados (Cartwright et al., 2016; Cummins et al., 2019; Watters et al., 2023).

Para além de seu uso como instrumento de avaliação, o FAST também se presta a procedimentos de treino relacional voltados ao aumento da flexibilidade no controle de estímulos. Embora o FAST ainda não tenha sido amplamente empregado como intervenção, pesquisas relacionadas utilizando o IAT têm mostrado resultados promissores. Em um estudo de grande escala, Lai et al. (2014) testaram dezessete metodologias diferentes de intervenção e constataram que a exposição repetida a exemplares inconsistentes com estereótipos no IAT esteve entre as mais eficazes na redução do viés racial implícito. O FAST é particularmente adequado a esforços semelhantes de redução de respostas relacionais racialmente enviesadas persistentes, por meio de seu reforço diferencial, tentativa a tentativa, de respostas relacionais flexíveis e sensíveis ao contexto. Tais aplicações alinham-se a modelos baseados em processos,

incluindo a Terapia de Aceitação e Compromisso (ACT), que conceitualizam, por exemplo, a defusão cognitiva como um meio de aumentar a flexibilidade comportamental e reduzir a dominância de regras verbais rígidas sobre o comportamento manifesto (Assaz et al., 2018; Assaz et al., 2023; Farmer et al., 2023; Hayes et al., 2004; Lillis & Hayes, 2007).

Apesar desse potencial conceitual, nenhum estudo anterior testou se um treino baseado no FAST, aplicado antes de uma tarefa relevante para o viés, pode reduzir respostas comportamentais racialmente enviesadas em contextos ecologicamente válidos e sob pressão de tempo. A Tarefa do Dilema do Policial (*Police Officer Dilemma Task – PODT*) é uma das medidas, amplamente utilizada para avaliar a tomada de decisão racialmente enviesada sob pressão (Correll et al., 2002). Nesse jogo digital de tiro em primeira pessoa, os participantes decidem, sob pressão temporal, se devem “atirar” ou “não atirar” em alvos armados ou desarmados brancos ou negros. Participantes brancos consistentemente apresentam taxas mais elevadas de falsos alarmes para alvos negros desarmados e decisões corretas de disparo mais rápidas para alvos negros armados, um padrão que reflete controle de estímulos racialmente enviesado (Payne & Correll, 2020). Essa medida emprega indicadores de tempo de resposta e acurácia, bem como métricas da Teoria da Detecção de Sinais, que podem ser compreendidas funcionalmente como índices derivados de controle contextual (ver Critchfield, 2002; Critchfield & Perone, 1993).

O presente estudo investiga se diferentes configurações de treino relacional baseadas no FAST podem reduzir o viés racial conforme mensurado pela PODT. Especificamente, os participantes foram alocados aleatoriamente em quatro

grupos: (a) grupo de treino de flexibilidade relacional (FAST completo com bloco consistente e inconsistente alternados), (b) grupo com apenas bloco de treino inconsistente com estereótipos, (c) grupo com apenas bloco de treino consistente com estereótipos e (d) grupo controle (sem intervenção). Os efeitos dessas diferentes configurações de intervenção foram avaliados por meio da PODT (medidas de acurácia, tempo, critério de decisão e discriminabilidade), da Escala de Racismo Moderno Modificada (MMRS, *Modified Modern Racism Scale*) e de avaliações de valência de faces novas negras e brancas por meio de uma Escala tipo Likert. Nesse estudo, esperava-se que desempenhos racialmente enviesados fossem menos pronunciados, em todas as medidas, nas condições de treino de flexibilidade e de treino inconsistente com estereótipos, em comparação com a condição controle e a condição de treino consistente com estereótipos.

## **Método**

### **Participantes e Delineamento**

Um total de 160 adultos brancos foi recrutado, dos quais 44 foram excluídos com base em critérios de qualidade dos dados (ver Análise de Dados), resultando em uma amostra final de 116 participantes ( $M = 23,13$  anos,  $DP = 3,18$ ; 67 mulheres). Todos os participantes foram remunerados, seja por meio de créditos em disciplina, como participantes de um curso de graduação em Psicologia, seja por meio de pagamento no valor aproximado de €6 pelo tempo despendido. Cinquenta participantes eram estudantes de graduação que receberam créditos acadêmicos, enquanto sessenta e sete participantes remunerados foram recrutados por meio da plataforma *Prolific*, especializada no recrutamento de

participantes para pesquisa. Todos os participantes eram residentes na República da Irlanda.

Os participantes foram alocados aleatoriamente a uma das quatro condições pelo software Inquisit, utilizado para controlar todas as fases do estudo. A condição de flexibilidade relacional (C1, n = 30) envolveu exposição alternada aos blocos de treino consistentes e inconsistentes com estereótipos do FAST (isto é, o procedimento FAST padrão). A Condição 2 (C2, n = 30) envolveu exposição repetida apenas ao bloco de treino inconsistente com estereótipos, enquanto a terceira condição (C3, n = 30) envolveu exposição repetida apenas ao bloco de treino consistente com estereótipos. Por fim, a quarta condição (C4) atuou como grupo controle, envolvendo exposição apenas às avaliações pós-intervenção de viés racial (C4, n = 26).

Uma análise de poder de sensibilidade realizada no software G\*Power 3.1 (Faul et al., 2007) indicou que o tamanho final da amostra (N = 116) forneceu 90% de poder estatístico ( $\alpha = .05$ ) para detectar um tamanho de efeito mínimo de  $f = 0,15$  para a interação intra-entre sujeitos em uma ANOVA mista.

## **Materiais**

### ***Function Acquisition Speed Test (FAST – Versão Treino)***

Três tipos de intervenção foram ministrados utilizando FAST na versão treino, as quais consistiram em: (a) treino com FAST completo, composta por dois blocos, consistente e inconsistente, administrada três vezes em sucessão (isto é, seis blocos; C1); ou (b) apenas um dos dois blocos do procedimento FAST padrão, administrado seis vezes em sucessão (isto é, C2 com bloco inconsistente

e C3 com bloco consistente). Todas as intervenções foram administradas on-line por meio do software Inquisit (millisecond.com).

Cada bloco de treino consistiu em 50 tentativas. O procedimento FAST padrão foi composto por dois blocos desse tipo: um estabelecendo duas classes de resposta funcional consistentes com um estereótipo racial presumido (isto é, Negro–Negativo e Branco–Positivo) e outro consistindo em 50 tentativas de treino que estabeleciam duas classes de resposta funcional inconsistentes com o estereótipo racial (isto é, Negro–Positivo e Branco–Negativo). A ordem dos blocos em cada uma das três administrações do FAST completo foi randomizada pelo software.

Cada tentativa de treino envolveu a apresentação de um único estímulo, pertencente a um de quatro conjuntos: um conjunto de quatro faces negras, um conjunto de quatro faces brancas, um conjunto de quatro palavras de atributos positivos (isto é, *Good, Love, Pleasure, Happy*, traduzidas para português, Bom, Amor, Prazer e Alegria) e um conjunto de quatro palavras de atributos negativos (isto é, *Bad, Filthy, Rotten, Evil*, traduzidas para o português Mau, Imundo, Podre, Maligno). Metade das imagens faciais negras e brancas era composta por faces masculinas e metade por faces femininas. Todos os estímulos foram selecionados a partir do conjunto original utilizado por Nosek et al. (2002) para o IAT de viés racial, disponível em [projectimplicit.com](http://projectimplicit.com). As apresentações dos estímulos ocorreram em ordem quase aleatória, com o software selecionando um estímulo de cada categoria a cada ciclo de quatro tentativas.

As contingências de reforçamento foram programadas para estabelecer, em cada bloco, classes de resposta funcional consistentes ou inconsistentes com

estereótipos. Isso foi feito por meio de instruções iniciais informando aos participantes que estivessem preparados para pressionar as teclas Z ou M do teclado do computador em resposta a cada estímulo apresentado na tela. As respostas corretas foram, a partir de então, guiadas apenas por feedback de resposta (isto é, as palavras “CORRECT” ou “WRONG”, exibidas em fonte verde ou vermelha, respectivamente, por 0,5 segundo). Um intervalo entre tentativas (ITI) de 500 ms, com a exibição de uma tela branca, precedeu a apresentação de cada estímulo. As palavras foram apresentadas em fonte tamanho 32, enquanto as imagens foram exibidas no tamanho de 250 × 250 pixels. A ausência de resposta dentro de uma janela de resposta de 3000 milissegundos resultou em feedback negativo, no registro da resposta como incorreta e no registro de um tempo de resposta de 3000 milissegundos (3 segundos) para essa tentativa. A contingência de reforçamento assemelhou-se a um esquema FR1 com *tempo limite*<sup>13</sup> de 3 segundos.

Durante as condições C2 e C3, o mesmo bloco de treino (isto é, inconsistente ou consistente com estereótipos, respectivamente) foi apresentado seis vezes em sucessão, com instruções genéricas sendo reapresentadas na tela no início de cada bloco de 50 tentativas.

Para avaliar o desempenho no FAST, foi adotada uma abordagem de pontuação baseada nas recomendações de Cummins e Roche (2020) e no arcabouço teórico delineado por Watters et al. (2023). Especificamente, para indexar diferenciais de fluência de resposta entre os blocos do FAST dentro da

---

<sup>13</sup> Cada resposta correta era seguida de feedback imediato, desde que emitida em até 3 segundos, favorecendo respostas rápidas e sob pressão temporal.

condição C1 (um indicador de viés de resposta), calculou-se a diferença entre as taxas de respostas corretas por minuto (CRPM) e incorretas por minuto (IRPM), resultando em um escore de diferencial de taxa de resposta (*response-rate differential*, RRD) para cada bloco. O escore geral do FAST foi então calculado como um diferencial de fluência de taxa (*rate-fluency differential*, RFD) entre os blocos, subtraindo-se o RRD do bloco inconsistente do RRD do bloco consistente. Esse procedimento gera um escore positivo quando as trajetórias de aprendizagem são caracterizadas por maior fluência nas classes de resposta funcional esperadas e estabelecidas pré-experimentalmente, caracterizadas como estereótipos antinegros.

### ***Tarefa do Dilema do Policial***

Foi utilizada uma versão pública do *Dilema do Policial* implementada no software Inquisit (Katja Borchert / *Millisecond.com*). Todas as imagens e parâmetros empregados foram idênticos aos utilizados na versão original da tarefa (Correll et al., 2002).

O jogo Dilema do Policial apresenta uma série de cenas visuais de fundo sobre as quais são sobrepostas imagens-alvo de indivíduos segurando diferentes objetos. Vinte cenários de fundo distintos, que variam de ambientes urbanos, como estações de trem, calçadas e prédios comerciais, a ambientes naturais, como parques, são utilizados em cada tentativa para aumentar a validade ecológica da tarefa. As imagens-alvo consistem em 40 personagens negros e 40 personagens brancos. Os personagens aparecem em primeiro plano em uma de quatro posturas: em pé ou ajoelhados, com objetos erguidos ou abaixados. Metade dos personagens segura uma arma de fogo (revólver prateado ou pistola

preta), enquanto a outra metade segura um objeto não letal (lata de alumínio, câmera, telefone celular ou carteira).

O Dilema do Policial inclui um bloco de treino com 80 tentativas, destinado à familiarização com o procedimento. O bloco de teste inclui 160 tentativas, sendo 40 com personagens negros armados, 40 com personagens negros desarmados, 40 com personagens brancos armados e 40 com personagens brancos desarmados. Cada tentativa inicia-se com a apresentação de uma cruz de fixação, seguida por uma série de imagens de fundo apresentadas rapidamente e de forma alternada, sem a presença de qualquer personagem-alvo sobreposto. O número de imagens de fundo varia aleatoriamente de 1 a 4, e a duração de cada imagem varia aleatoriamente entre 500 e 1000 milissegundos. A imagem de fundo final permanece na tela e recebe a sobreposição de um personagem-alvo selecionado aleatoriamente. Os participantes são instruídos, desde o início, a responder prontamente ao detectar a imagem-alvo pressionando uma de duas teclas do teclado. Especificamente, devem pressionar a tecla "A" para atirar no alvo e a tecla "L" para não atirar.

Foi implementada uma contingência diferencial de reforçamento, conforme o procedimento padrão, na qual pressionar a tecla "A - Atirar" na presença de um alvo armado resultava em feedback positivo, "bom tiro" (*good shot*) e no ganho de 10 pontos, enquanto pressionar "A" na presença de um alvo desarmado resultava na mensagem "você atirou em um cara legal" (*you shot a good guy*), e resultava na perda de 20 pontos. Pressionar "L – Não Atirar" na presença de um alvo desarmado resultava em feedback positivo "sábua escolha" ("wise choice"), e no ganho de 5 pontos; já na presença de um alvo armado, era exibida a mensagem

"você morreu" ("you are dead"), resultando na perda de 40 pontos. Essas contingências de reforçamento foram estruturadas para reproduzir cenários do mundo real enfrentados por policiais, nos quais o objetivo central é evitar a identificação incorreta de um alvo armado e hostil, situação que pode resultar na morte do agente (ver Correll et al., 2002). Para minimizar omissões de resposta, participantes que não pressionavam nenhuma das teclas ("A" ou "L") dentro de um intervalo de 850 milissegundos sofriam uma dedução de 10 pontos e receberam o feedback "muito lento" ("too slow"). É importante destacar que esses pontos são totalmente arbitrários e não têm qualquer consequência sobre as métricas de desempenho utilizadas pelos pesquisadores.

As métricas de interesse no Dilema do Policial incluem: (1) acurácia, refletindo a capacidade do participante de identificar corretamente alvos armados e desarmados entre as diferentes raças; (2) tempo de reação para respostas corretas, indicando a velocidade de engajamento diante de indivíduos armados e desarmados; (3) falsos alarmes, definidos como disparos incorretos contra alvos desarmados; (4) *d-prime* ( $d'$ ), um índice da Teoria da Detecção de Sinais que captura a discriminabilidade entre alvos armados e desarmados; e (5) critério ( $c$ ), que reflete viés de resposta, indicando uma tendência geral a decisões de disparo mais conservadoras ou mais liberais. Os índices  $d'$  e  $c$  foram calculados separadamente para cada uma das duas raças de alvo. Essas métricas são fundamentadas na Teoria da Detecção de Sinais e amplamente utilizadas em estudos com o Dilema do Policial (por exemplo, Correll et al., 2002, 2006, 2007, 2014). No presente estudo, a acurácia foi tratada como a métrica primária de

desfecho. Análises complementares foram conduzidas para os tempos de reação, falsos alarmes,  $d'$  e  $c$ .

### **Escala de Racismo Moderno Modificada**

A Escala de Racismo Moderno Modificada (MMRS), originalmente desenvolvida por McConahay (1986), foi adaptada ao contexto cultural irlandês seguindo as recomendações de Barnes-Holmes, Murphy et al. (2010) e tem sido empregada em outros estudos, incluindo Roche et al. (2025). A MMRS consiste em seis afirmações, cada uma avaliada em uma escala de 5 pontos que varia de “Discordo totalmente” a “Concordo totalmente”. Diferentemente de medidas tradicionais, a MMRS foi concebida para capturar expressões sutis de racismo. Por exemplo, uma das afirmações é: “Nos últimos anos, o governo e a mídia têm demonstrado mais respeito às pessoas negras do que elas merecem”.

As respostas na MMRS são pontuadas em uma escala de  $-2$  a  $+2$ , com escore mínimo de  $-12$  e máximo de  $+12$  (ver uso exemplificado em Roche et al., 2025). Escores positivos indicam maior endosso de atitudes racistas contemporâneas (viés *antinegro*), escores próximos de zero refletem ausência de viés, e escores negativos baixa concordância com atitudes racistas contemporâneas.

### **Escala Likert**

A avaliação explícita de faces negras e brancas foi registrada por meio de uma escala Likert de 7 pontos, utilizada para mensurar a valência de diferentes imagens faciais. Uma avaliação de 1 representava “desagradável” e uma avaliação de 7 representava “agradável”. Os participantes foram apresentados sequencialmente a um total de oito faces.

Os estímulos faciais foram obtidos do banco de dados *300 Faces in-the-Wild* (300-W; Sagonas et al., 2016), um conjunto padronizado que contém fotografias faciais de alta qualidade e diversidade racial, coletadas em condições naturalísticas, permitindo uso experimental controlado e, ao mesmo tempo, ecologicamente válido. Os mesmos estímulos já foram utilizados em estudos anteriores sobre viés racial (Mizael et al., 2016, 2021; Passarelli, Roche et al., 2024; Passarelli, Mizael et al., 2025). O conjunto selecionado foi composto por quatro faces negras (duas femininas e duas masculinas) e quatro faces brancas (duas femininas e duas masculinas). A ordem de apresentação das faces na tela para avaliação foi randomizada entre as tentativas. Essas imagens eram diferentes daquelas utilizadas nas intervenções com o FAST e foram mantidas distintas para permitir inferências sobre a generalização do efeito da intervenção para a avaliação de imagens faciais novas representativas de diferentes grupos raciais.

### **Procedimento**

O experimento foi administrado por meio do software Inquisit, disponibilizado on-line e acessível por qualquer navegador da web. Os participantes remunerados realizaram o estudo remotamente. Os participantes que receberam créditos acadêmicos acessaram o estudo em computadores fornecidos presencialmente nas instalações de pesquisa da universidade. Os pequenos cubículos de pesquisa nos quais participaram tinham aproximadamente 4 m<sup>2</sup>.

Para todos os participantes, informações sobre o estudo (sem revelar as hipóteses experimentais), um termo de consentimento e um breve questionário demográfico foram apresentados antes da aplicação da intervenção FAST. A

alocação às condições foi randomizada pelo software no início de cada intervenção, sem contrabalanceamento entre condições. No grupo de flexibilidade (C1), os participantes foram expostos a blocos alternados de aprendizagem consistente e inconsistente com estereótipos, com cada par repetido três vezes (isto é, 100 tentativas por administração). Cada bloco foi precedido pelas mesmas instruções. Na condição C2 (inconsistente com estereótipos), o bloco inconsistente do FAST, composto por 50 tentativas, foi repetido seis vezes (total de 300 tentativas), com cada bloco precedido pelas mesmas instruções sobre como realizar a tarefa. De modo semelhante, na condição C3 (consistente com estereótipos), o bloco consistente do FAST foi repetido seis vezes, sempre iniciado com as mesmas instruções. Por fim, os participantes da condição C4 não receberam intervenção, mas realizaram as avaliações pós-intervenção, assim como os demais grupos.

A primeira avaliação pós-intervenção foi a PODT, que se iniciou com um bloco de prática de 80 tentativas e, em seguida, prosseguiu para um bloco de teste de 160 tentativas. Subsequentemente, os participantes responderam à MMRS e avaliaram faces negras e brancas em uma escala Likert. Após a conclusão das medidas, os participantes receberam um *debriefing* apresentado na tela.

### **Ética e Transparência da Pesquisa**

O presente estudo foi aprovado pelo Subcomitê de Ética em Pesquisa em Ciências Biomédicas e da Vida da Maynooth University (ID: 37220) e seguiu o Código de Ética da Psychological Society of Ireland, bem como as regulamentações da União Europeia relativas à Proteção Geral de Dados (GDPR).

## **Análise de Dados**

Antes das análises, os dados foram examinados quanto à qualidade e à conformidade com as tarefas, com exclusões baseadas em: questionários pós-teste incompletos ( $n = 17$ ), respostas invariantes durante as tentativas do FAST ( $n = 6$ ), blocos de treino FAST incompletos ( $n = 13$ ) ou não conclusão da tarefa PODT ( $n = 8$ ).

Para avaliar a fluência de resposta no desempenho do FAST, foram empregados os índices RRD e RFD, conforme sugerido por Cummins e Roche (2020; ver seção FAST acima). Essa abordagem permitiu calcular índices de fluência para cada bloco repetido do FAST e, para os participantes da condição de flexibilidade, derivar um escore geral de diferencial de fluência agregado ao longo de três pares de blocos (isto é, três escores sucessivos de viés). Gráficos de dispersão foram utilizados para representar visualmente as mudanças nos escores de fluência ao longo dos blocos sucessivos do FAST (C2 e C3) e os escores de viés do FAST baseados nos diferenciais de fluência entre blocos (apenas C1). Um coeficiente linear de aprendizagem foi calculado para cada participante por meio de regressão dos escores de fluência sobre blocos de treino idênticos sucessivos (RRD; C2 e C3) ou exposições completas ao FAST (RFD; C1). Inclinações positivas para C2 e C3 indicariam aumentos na fluência de resposta ao longo dos blocos repetidos. Para C1, uma inclinação negativa indicaria redução do diferencial de fluência entre pares de blocos ao longo das três administrações sucessivas do FAST.

Na PODT, as variáveis dependentes primárias foram a acurácia das respostas e o tempo de reação para respostas corretas, analisados por meio de

uma ANOVA mista com delineamento intra–entre sujeitos 2 (Tipo de Objeto: Arma vs. Sem Arma) × 2 (Raça do Alvo: Negro vs. Branco) × 4 (Condição: Inconsistente, Consistente, Flexibilidade, Controle). Além disso, foram analisados falsos alarmes, *d-prime* ( $d'$ ) e critério ( $c$ ) por meio de uma ANOVA mista separada 2 (Raça do Alvo) × 4 (Condição), com medidas repetidas em raça e fator entre sujeitos para condição. Falsos alarmes foram definidos como disparos contra alvos desarmados. O *d-prime* foi calculado como:  $d' = Z(\text{taxa de acertos}) - Z(\text{taxa de falsos alarmes})$ , em que os valores  $Z$  foram obtidos convertendo as taxas de acertos e falsos alarmes em escores normais padronizados por meio de consulta reversa à distribuição normal, permitindo expressar o desempenho em uma escala comum. Esse índice representa a discriminabilidade entre armas e objetos não letais, calculada separadamente para alvos negros e brancos. O critério foi calculado como:  $c = -0,5 \times [Z(\text{taxa de acertos}) + Z(\text{taxa de falsos alarmes})]$ . Esse índice representa o viés de resposta, com valores mais altos indicando um limiar mais conservador para atirar e valores mais baixos indicando um viés mais liberal de disparo, calculado separadamente para alvos negros e brancos.

As avaliações em escala Likert para faces negras e brancas foram analisadas por meio de uma ANOVA mista 2 (Raça do Alvo) × 4 (Condição). Por fim, os escores totais da MMRS foram analisados por meio de uma ANOVA unifatorial entre condições. Correlações de Pearson foram calculadas para examinar a covariância entre os escores individuais de fluência do FAST por bloco (RRD; Condições 2 e 3) e as medidas de desfecho da PODT. Adicionalmente, para fins exploratórios, foram examinadas correlações entre os aumentos de fluência ao longo dos blocos individuais de treino (inclinação de aprendizagem) e

múltiplas medidas de desfecho (MMRS, Escala Likert, Acurácia, Tempo de Reação, critério, *d-prime*). No caso da condição C1, foram examinadas correlações entre reduções nos escores de viés ao longo das administrações completas do FAST (escores RFD) e as mesmas medidas de desfecho. Por fim, correlações entre as avaliações de faces e os escores da *MMRS* com as métricas da PODT permitiram quantificar as relações entre medidas explícitas e implícitas.

## Resultados

### Desempenho nas Intervenções FAST

Na condição de Flexibilidade (C1), os blocos do FAST foram administrados três vezes cada, em pares apresentados em ordem aleatória. Para os blocos consistentes, o diferencial de taxa de resposta médio (RRD) na primeira apresentação foi de 20,75 ( $DP = 2,69$ ), aumentando para 23,85 ( $DP = 2,23$ ) na terceira apresentação. Nos blocos inconsistentes, o RRD apresentou média inicial de 17,50 ( $DP = 3,52$ ), aumentando para 22,25 ( $DP = 4,43$ ) na terceira exposição. Apesar do aumento da fluência de resposta ao longo dos blocos sucessivos no FAST completo, observou-se uma diminuição da fluência diferencial de resposta (RFD) entre pares sucessivos de blocos, com os escores diminuindo de 3,25 ( $DP = 2,25$ ) na primeira apresentação para 1,60 ( $DP = 2,38$ ) na terceira. Esse padrão indica uma redução do viés de resposta e um aumento da flexibilidade de resposta ao longo das administrações completas do FAST.

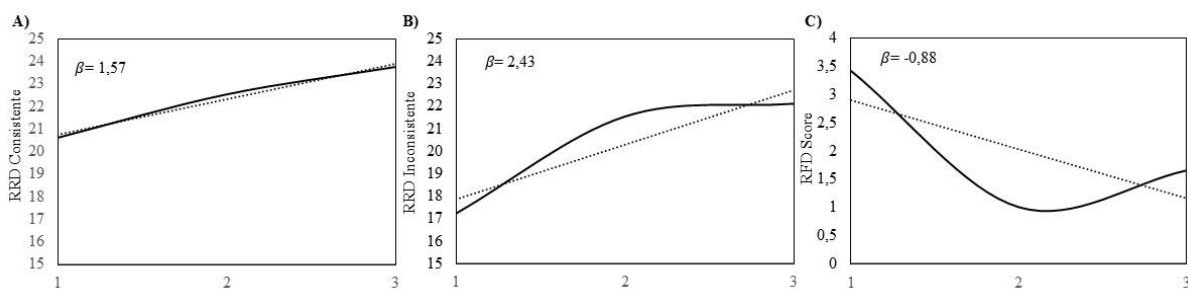
Para examinar esses efeitos de forma mais detalhada na condição C1, foram calculadas inclinações de aquisição ( $\beta$ ) para os escores de RRD ao longo das três apresentações dos blocos consistentes, para os escores de RRD ao longo das três apresentações dos blocos inconsistentes, e para três escores

sucessivos de RFD derivados de pares sucessivos de blocos consistentes e inconsistentes. Foram observadas inclinações médias positivas de mudança no RRD tanto para os blocos consistentes ( $\beta$  média = 1,57, DP = 2,27, IC 95% [0,75, 2,58]) quanto para os blocos inconsistentes ( $\beta$  média = 2,43, DP = 4,57, IC 95% [0,98, 3,75]), refletindo aumento de fluência ao longo das repetições. Em contraste, e conforme esperado, o RFD apresentou inclinação média negativa (-0,88, DP = 3,68, IC 95% [-3,75, 0,88]), indicando uma redução do diferencial de fluência entre os blocos consistentes e inconsistentes (isto é, redução do viés).

Os painéis A e B da Figura 1 apresentam as mudanças de fluência (RRD) ao longo das apresentações repetidas dos blocos na condição de Flexibilidade (C1). As linhas contínuas representam as médias observadas e as linhas pontilhadas representam as funções de regressão linear ajustadas. As inclinações de regressão ( $\beta$ ) são exibidas em cada painel, indexando as mudanças na fluência ao longo das repetições para os blocos consistentes (Painel A), inconsistentes (Painel B) e para o diferencial de fluência de resposta (RFD; Painel C).

### Figura 1

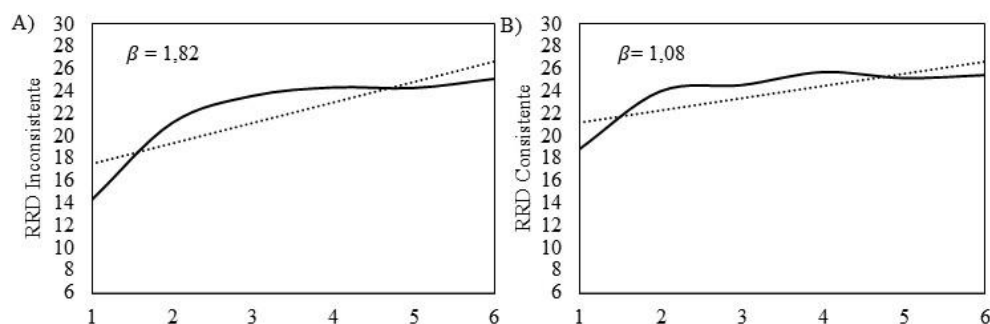
*Fluência de resposta (RRD) e diferencial de fluência de resposta (RFD) na condição de flexibilidade (C1)*



Para a condição inconsistente com estereótipos (C2), a fluência média (RRD) aumentou ao longo das seis apresentações sucessivas dos blocos, de  $M = 14,32$  ( $DP = 9,78$ ) na primeira exposição para  $M = 25,35$  ( $DP = 3,52$ ) na sexta. A inclinação de aquisição média entre os participantes foi positiva ( $1,82$ ,  $DP = 1,67$ ,  $IC\ 95\% [0,50, 2,25]$ ), indicando aumento da fluência com a repetição da tarefa. Um padrão comparável foi observado na condição consistente com estereótipos (C3). A fluência média aumentou de  $M = 18,82$  ( $DP = 7,84$ ) na primeira exposição para  $M = 25,44$  ( $DP = 5,83$ ) na sexta. A inclinação média de aprendizagem também foi positiva ( $1,57$ ,  $DP = 1,25$ ,  $IC\ 95\% [0,65, 1,58]$ ), refletindo aumento da fluência ao longo da intervenção. A Figura 2 apresenta a fluência média de resposta (RRD) ao longo das exposições aos blocos para a condição inconsistente com estereótipos (Painel A: C2) e para a condição consistente com estereótipos (Painel B: C3). As linhas contínuas representam o desempenho médio observado e as linhas pontilhadas representam as funções de regressão linear ajustadas, cujas inclinações ( $\beta$ ) indicam a taxa de mudança ao longo da intervenção. Nas três condições experimentais observou-se aumento na fluência de desempenho no FAST ao longo das sessões. A primeira condição diferenciou-se das demais por apresentar, adicionalmente, redução do viés racial, sugerindo que o aumento de fluência por si só não é suficiente para produzir mudança no padrão relacional e que a configuração específica das contingências de treino é determinante para os efeitos sobre o viés.

## Figura 2

Fluência de resposta ao longo de seis exposições aos blocos nas condições C2 (inconsistente com estereótipos; Painel A) e C3 (consistente com estereótipos; Painel B).



## Dilema do Policial (PODT)

### Acurácia

Foi conduzida uma ANOVA mista com delineamento 2 (Raça: Negro, Branco)  $\times$  2 (Tipo de Objeto: Arma, Sem Arma), como fatores intra-sujeitos,  $\times$  4 (Condição: Inconsistente, Consistente, Flexibilidade, Controle), como fator entre sujeitos, para avaliar a acurácia das respostas. Embora não tenham sido observados efeitos principais significativos para tipo de objeto ou raça, emergiu uma interação significativa de três fatores entre condição, tipo de objeto e raça,  $F(3, 112) = 4,95, p = .003, \eta^2 = .12$ . Comparações pareadas *post hoc* com ajuste de Bonferroni indicaram que, para alvos brancos desarmados, os participantes da condição controle ( $M = .95, DP = .05$ ) foram significativamente mais acurados em não atirar do que aqueles da condição de flexibilidade ( $M = .88, DP = .06$ ),  $p = .004, IC 95\% [-.14, -.02]$ . Nenhuma outra comparação pareada entre condições alcançou significância estatística.

Comparações pareadas com ajuste de Bonferroni dentro de cada condição foram conduzidas para examinar as interações entre raça (Negro vs. Branco) e tipo de objeto (Arma vs. Sem Arma). Nas condições de flexibilidade e inconsistente com estereótipos, não foram observadas diferenças significativas (todos os valores de  $p > .05$ ) entre alvos negros e brancos, tanto para decisões de “atirar” envolvendo alvos armados quanto para decisões de “não atirar” envolvendo alvos desarmados, indicando ausência de evidências de viés racial nessas condições.

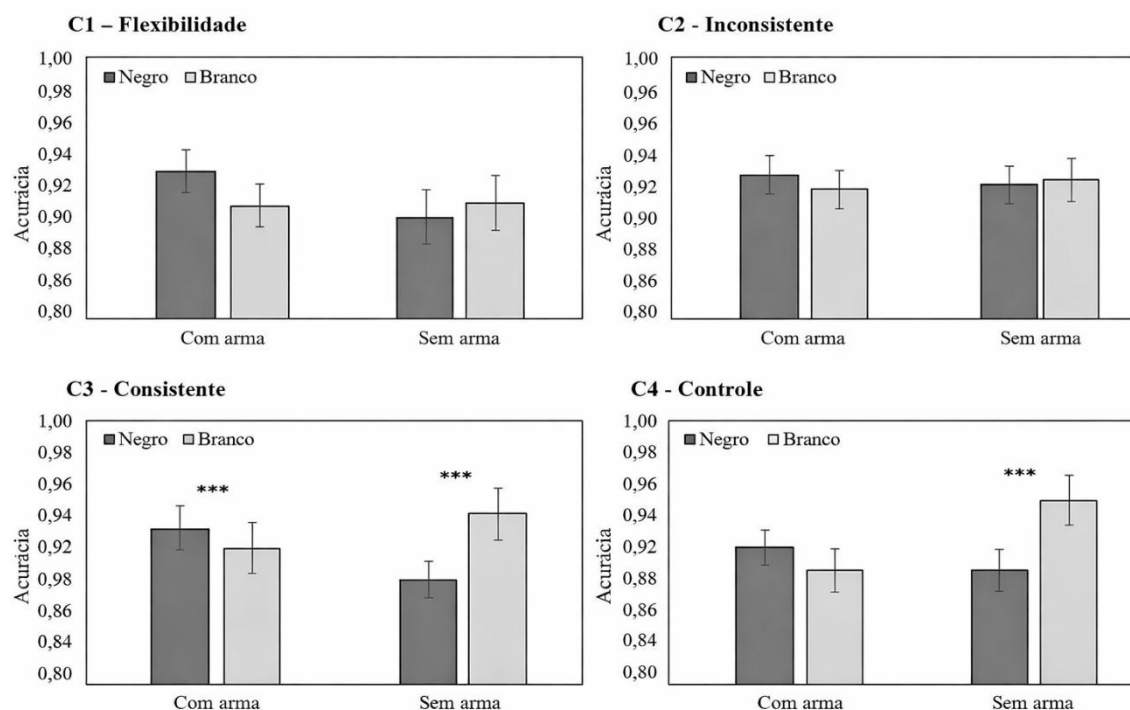
Em contraste, diferenças significativas foram observadas na condição consistente com estereótipos. Os participantes foram mais acurados ao atirar em alvos negros armados do que em alvos brancos armados (C3;  $DM = .05$ ,  $EP = .02$ ,  $p = .002$ , IC 95% [.02, .09]), um padrão consistente com viés antinegro. Além disso, os participantes apresentaram maior acurácia em decisões de “não atirar” para alvos brancos desarmados em comparação com alvos negros desarmados tanto na condição consistente com estereótipos (C3;  $DM = -.08$ ,  $EP = .015$ ,  $p < .001$ , IC 95% [-0.111, -0.052]) quanto na condição controle (C4;  $DM = -.067$ ,  $EP = .02$ ,  $p < .001$ , IC 95% [-0.10, -0.04]), refletindo um viés pró-branco nas decisões de abstenção de disparo. Esses resultados indicam que a condição consistente com estereótipos não apenas deixou o viés inalterado, mas produziu aumento no padrão de respostas enviesadas, sugerindo que o treino de relações congruentes com o viés socialmente dominante pode reforçar e ampliar esse padrão relacional.

A Figura 3 apresenta as taxas de acurácia para decisões de “atirar” e “não atirar” em função do tipo de objeto, da raça do alvo e da condição experimental.

Diferenças pareadas significativas são indicadas por asteriscos (\*\*\*) , representando efeitos ao nível de  $p < .001$ .

### Figura 3

*Acurácia das decisões em relação aos tipos de objeto para alvos negros e brancos, ao longo das condições experimentais*



*Nota.* As barras pretas e cinzas correspondem às respostas a alvos negros e brancos, respectivamente. As barras de erro representam  $\pm 1$  erro-padrão da média (EP).

### Tempo de Reação

O tempo de reação foi analisado por meio de uma ANOVA mista com delineamento 2 (Raça: Negro, Branco)  $\times$  2 (Tipo de Objeto: Arma, Sem Arma) como fatores intra-sujeitos  $\times$  4 (Condição) como fator entre sujeitos. Observou-se um efeito principal significativo do tipo de objeto,  $F(1, 112) = 446,03$ ,  $p < .001$ ,  $\eta p^2 = .80$ , indicando respostas mais rápidas a alvos armados do que a alvos

desarmados. Comparações pareadas com ajuste de Bonferroni revelaram um efeito robusto do tipo de objeto sobre o tempo de reação. Especificamente, considerando todas as condições, os participantes responderam significativamente mais rápido quando os alvos estavam armados do que quando estavam desarmados,  $DM = -62,00$  ms,  $EP = 2,94$ ,  $p < .001$ , IC 95%  $[-67,82, -56,18]$ . Esse padrão foi consistente em todas as condições experimentais, com respostas mais rápidas a alvos armados em comparação aos desarmados.

Também emergiu uma interação significativa de dois fatores entre Tipo de Objeto  $\times$  Raça,  $F(1, 112) = 9,26$ ,  $p = .003$ ,  $\eta p^2 = .08$ . Comparações pareadas com ajuste de Bonferroni examinando essa interação indicaram diferenças significativas nos tempos de reação entre tipos de objeto dentro de cada categoria racial. Para alvos negros, os participantes responderam significativamente mais rápido a alvos armados do que a alvos desarmados,  $DM = -69,11$  ms,  $EP = 4,02$ ,  $p < .001$ , IC 95%  $[-77,07, -61,15]$ . Um padrão semelhante foi observado para alvos brancos, com respostas mais rápidas a alvos armados em comparação a alvos desarmados,  $DM = -54,89$  ms,  $EP = 3,47$ ,  $p < .001$ , IC 95%  $[-61,76, -48,03]$ .

Comparações dos tempos de reação entre alvos de diferentes raças dentro de cada tipo de objeto indicaram ainda que os participantes responderam corretamente de forma significativamente mais rápida a alvos negros armados do que a alvos brancos armados,  $DM = -8,68$  ms,  $EP = 3,23$ ,  $p = .008$ , IC 95%  $[-15,08, -2,29]$ . Em contraste, para alvos desarmados, as respostas corretas foram significativamente mais rápidas para alvos brancos do que para alvos negros,  $DM = 5,53$  ms,  $EP = 2,60$ ,  $p = .035$ , IC 95%  $[-3,39, 10,68]$ .

### **Critério de Decisão (c)**

O critério de decisão (c) é um índice de conservadorismo da resposta. Ele foi analisado por meio de uma ANOVA mista com delineamento 4 (Condição: inconsistente com estereótipos, consistente com estereótipos, flexibilidade, controle)  $\times$  2 (Raça: alvos negros vs. brancos), tendo Raça como fator intra-sujeitos e Condição como fator entre sujeitos. Valores mais elevados de critério indicam uma estratégia de decisão mais conservadora<sup>14</sup>.

A análise revelou um efeito principal significativo de Raça,  $F(1, 112) = 14,76$ ,  $p < .001$ ,  $\eta p^2 = .12$ , indicando que, de modo geral, os participantes foram mais conservadores ao responder a alvos brancos do que a alvos negros. Comparações pareadas com ajuste de Bonferroni confirmaram esse efeito, com valores de critério significativamente mais altos para alvos brancos em comparação com alvos negros ( $DM = .36$ ,  $EP = .09$ ,  $p < .001$ , IC 95% [.17, .54]).

De forma importante, foi observada uma interação significativa Raça  $\times$  Condição,  $F(3, 112) = 5,30$ ,  $p = .002$ ,  $\eta p^2 = .12$ , indicando que as diferenças raciais no critério de decisão variaram em função da condição experimental. Análises de efeitos simples com ajuste de Bonferroni mostraram que os participantes foram significativamente mais conservadores em relação a alvos brancos do que a alvos negros na condição consistente com estereótipos (C3;  $DM = .78$ ,  $EP = .18$ ,  $p < .001$ , IC 95% [.42, 1.14]) e na condição controle (C4;  $DM = .58$ ,  $EP = .19$ ,  $p = .003$ , IC 95% [.20, .97]). Não foram observadas diferenças

---

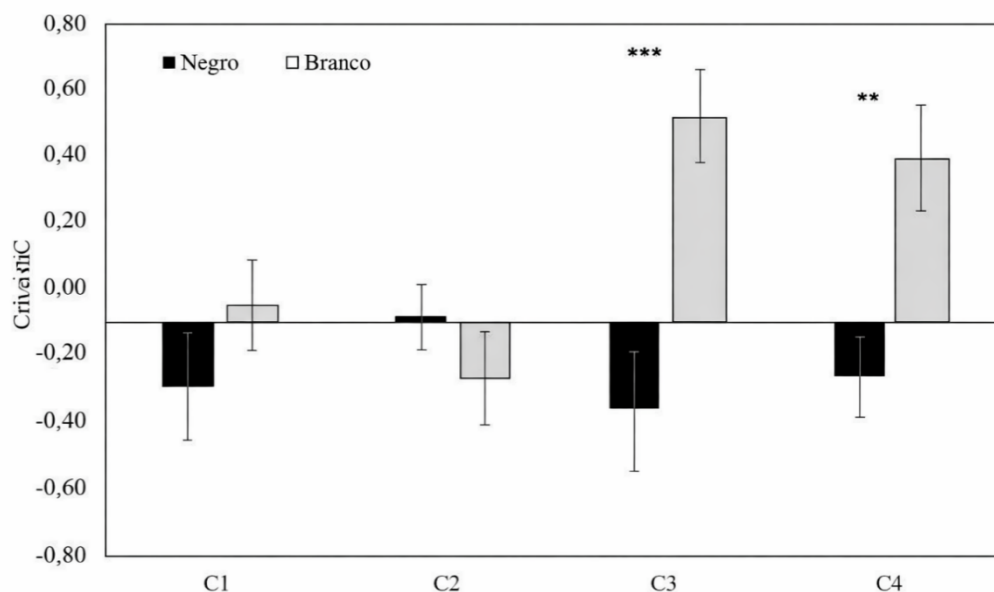
<sup>14</sup> Uma decisão mais conservadora corresponde à adoção de um critério de resposta mais rígido, no qual o participante evita atirar diante de incerteza. Esse padrão se manifesta por redução de falsos alarmes (atirar em alvos desarmados), mesmo ao custo de perda de acertos (deixar de atirar em alvos armados), refletindo mudança no critério decisório, e não na sensibilidade discriminativa ( $d'$ ).

raciais significativas nos escores de critério nas condições inconsistente com estereótipos (C2;  $p = .34$ ) ou de flexibilidade (C1;  $p = .21$ ).

A Figura 4 apresenta os valores médios de critério ( $c$ ) para alvos negros e brancos ao longo das quatro condições experimentais. Os asteriscos indicam diferenças estatisticamente significativas entre alvos negros e brancos (\* $p < .05$ , \*\* $p < .01$ , \*\*\* $p < .001$ ; ajuste de Bonferroni).

#### Figura 4

*Escores médios de critério de decisão (c) para alvos negros e brancos ao longo das quatro condições experimentais*



*Nota.* Escores mais elevados indicam decisões de disparo mais conservadoras. As barras pretas e cinzas correspondem às respostas a alvos negros e brancos, respectivamente. As barras de erro representam  $\pm 1$  erro-padrão da média (EP).

#### **Sensibilidade (d-prime)**

A sensibilidade ( $d'$ ) foi analisada por meio de uma ANOVA mista com delineamento 2 (Raça: Negro, Branco)  $\times$  4 (Condição: Flexibilidade, Inconsistente,

Controle, Consistente). Observou-se um efeito principal significativo de Raça,  $F(1, 112) = 8,47, p = .004, \eta^2 = .07$ , indicando diferenças gerais na sensibilidade discriminativa entre alvos negros e brancos. Comparações pareadas com ajuste de Bonferroni indicaram que a sensibilidade ( $d'$ ) foi significativamente maior para alvos brancos do que para alvos negros, de modo geral ( $DM = 0,27, EP = 0,09, p = .004, IC\ 95\% [.09, .46]$ ). Em termos simples, os participantes foram mais precisos ao discriminar armas de objetos não letais quando o alvo era branco e mais propensos a cometer erros de discriminação quando o alvo era negro. Em termos analítico-comportamentais, a cor da pele exerceu controle primário sobre a resposta operante: ver uma pessoa branca funcionou como estímulo discriminativo que evocou respostas de observação adicionais para verificar a presença ou ausência de arma antes de responder, enquanto ver uma pessoa negra evocou diretamente a resposta de atirar, independentemente do objeto portado, indicando que o controle pela raça antecedeu e suprimiu o controle pelo estímulo relevante para a tarefa.

Foi observada uma interação significativa Raça  $\times$  Condição,  $F(3, 112) = 7,73, p < .001, \eta^2 = .17$ , indicando que as diferenças raciais na sensibilidade variaram em função da condição experimental. Não houve efeito principal de Condição,  $F(3, 112) = 0,22, p = .88$ . Análises de efeitos simples com ajuste de Bonferroni indicaram diferenças significativas de sensibilidade entre alvos negros e brancos tanto na condição Consistente quanto na condição Controle.

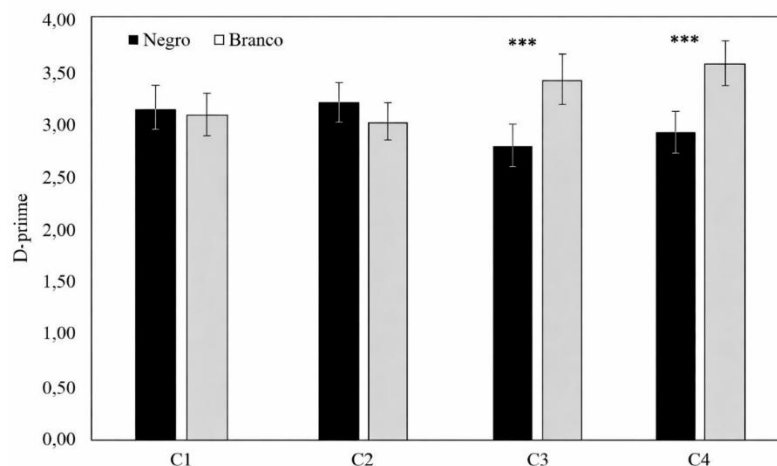
Na condição Consistente (C3), a sensibilidade foi maior para alvos brancos do que para alvos negros ( $DM = -.76, EP = 0,18, p < .001, IC\ 95\% [-1,13, -0,40]$ ). De forma semelhante, na condição Controle (C4), a sensibilidade também foi

maior para alvos brancos do que para alvos negros ( $DM = -0,67$ ,  $EP = 0,20$ ,  $p < .001$ , IC 95%  $[-1,06, -0,27]$ ). Não foram observadas diferenças significativas de sensibilidade entre alvos negros e brancos nas condições Inconsistente ou de Flexibilidade (ambos os  $ps > .05$ ).

A Figura 5 apresenta os valores médios de  $d'$  para alvos negros e brancos ao longo das quatro condições experimentais. Os asteriscos indicam diferenças estatisticamente significativas entre alvos negros e brancos ( $*p < .05$ ,  $**p < .01$ ,  $***p < .001$ ; ajuste de Bonferroni).

### Figura 5

*Escores de sensibilidade ( $d'$ ) para alvos negros e brancos ao longo das quatro condições experimentais*



*Nota.* Valores mais altos de  $d'$  refletem melhor discriminação entre alvos armados e desarmados. As barras pretas e cinzas representam, respectivamente, os escores de sensibilidade ( $d'$ ) para alvos negros e brancos. As barras de erro representam  $\pm 1$  erro-padrão da média (EP).

### Escala de Racismo Moderno Modificada (MMRS)

De modo geral, os participantes relataram explicitamente atitudes não racistas na Escala de Racismo Moderno Modificado (MMRS), uma vez que as

médias foram negativas em todas as condições experimentais (C1:  $M = -8,63$ ,  $DP = 4,76$ ; C2:  $M = -9,30$ ,  $DP = 3,25$ ; C3:  $M = -8,00$ ,  $DP = 4,14$ ; C4:  $M = -7,96$ ,  $DP = 4,80$ ). Uma ANOVA one-way não indicou diferenças significativas nos escores da MMRS entre as condições,  $F(3, 115) = 0,64$ ,  $p = .59$ ,  $\eta^2_p = .02$ .

### **Escala Likert**

Uma ANOVA mista com delineamento 2 (Raça: Negro, Branco)  $\times$  4 (Condição: flexibilidade, inconsistente com estereótipos, consistente com estereótipos, controle) revelou um efeito principal significativo de Raça,  $F(1, 112) = 6,91$ ,  $p = .01$ ,  $\eta^2_p = .06$ , nas avaliações explícitas de faces. Comparações pareadas com ajuste de Bonferroni indicaram que, de modo geral, as faces negras foram avaliadas mais positivamente do que as faces brancas ( $DM = .23$ ,  $EP = .09$ ,  $p = .01$ , IC 95% [0,06, 0,40]). A interação Raça  $\times$  Condição não foi significativa,  $F(3, 112) = 0,92$ ,  $p = .44$ ,  $\eta^2_p = .02$ , tampouco o efeito principal de Condição,  $F(3, 112) = 1,93$ ,  $p = .13$ ,  $\eta^2_p = .05$ .

### **Correlações entre as Medidas**

Uma análise de correlações entre as variáveis foi conduzida com fins exploratórios. Um pequeno número de correlações significativas foi observado. Especificamente, na condição de flexibilidade (C1), foi observada uma correlação positiva significativa entre a inclinação de aprendizagem no FAST (isto é, redução do viés indexada pelo escore RFD) e os tempos de reação quando os participantes corretamente se abstiveram de atirar em alvos brancos desarmados ( $r = .43$ ,  $p = .02$ ).

Na condição inconsistente com estereótipos (C2), escores mais baixos na MMRS (indicativos de viés pró-negro) correlacionaram-se com maior acurácia em decisões corretas de não atirar em alvos negros desarmados ( $r = -.36, p = .048$ ). De forma semelhante, escores mais baixos na MMRS estiveram associados a maior discriminabilidade ( $d'$ ) para alvos negros ( $r = -.41, p = .03$ ), sugerindo maior precisão na tomada de decisão à medida que o preconceito explícito diminuía.

Na condição consistente com estereótipos (C3), escores mais baixos na MMRS também se correlacionaram com maior acurácia em decisões de não atirar em alvos negros desarmados ( $r = -.36, p = .048$ ). Por fim, no grupo controle (C4), foi observada uma correlação negativa significativa entre as avaliações de valência de faces negras e os tempos de reação para disparos corretos em alvos brancos armados ( $r = -.40, p = .043$ ). Ou seja, avaliações mais positivas de faces negras estiveram associadas a menores tempos de reação ao atirar corretamente em alvos brancos armados.

### **Discussão**

O estudo explorou os efeitos de uma intervenção baseada no *Function Acquisition Speed Test* (FAST) no aumento da flexibilidade relacional e seus efeitos profiláticos consequenciais sobre o desempenho na *Police Officer's Dilemma Task* (PODT; Correll et al., 2002) e em medidas explícitas de viés. Quatro tipos de intervenção foram comparados: treino de flexibilidade relacional, treino inconsistente com estereótipos, treino consistente com estereótipos e um grupo controle sem intervenção. Todos os grupos de intervenção demonstraram efeitos positivos de aprendizagem ao longo dos blocos de intervenção, indicando aquisição de classes funcionais e engajamento na tarefa. O grupo de flexibilidade

apresentou uma redução no escore de diferencial de fluência de taxa (RFD) ao longo dos pares de blocos de intervenção, consistente com maior flexibilidade relacional.

No que se refere ao comportamento de disparo, participantes nas condições de flexibilidade e inconsistente com estereótipos não apresentaram vieses raciais na acurácia. Em contraste, participantes na condição consistente com estereótipos foram mais precisos ao atirar em alvos negros armados em comparação a alvos brancos armados, enquanto tanto o grupo consistente com estereótipos quanto o grupo controle foram mais precisos em não atirar em alvos brancos desarmados do que em alvos negros desarmados. Esses resultados são consistentes com estudos anteriores e tipicamente associados ao viés racial (Correll et al., 2002; Correll et al., 2006; Correll et al., 2007; Correll et al., 2014; Ross, 2015). Assim, parece que as intervenções de treino relacional de flexibilidade e de inconsistência com estereótipos modificaram efetivamente as respostas relacionais a alvos raciais em uma simulação de tiro em primeira pessoa sob pressão temporal que modela um dilema policial do mundo real. Em outras palavras, intervenções breves de treino relacional parecem ser capazes de “direcionar” vieses de resposta relacional.

As análises de tempo de reação revelaram uma interação significativa entre tipo de objeto e raça do alvo, com o conjunto de participantes, como um todo, respondendo mais rapidamente com decisões de disparo para alvos negros armados do que para alvos brancos armados, e sendo mais rápidos para decidir não atirar em alvos brancos desarmados do que em alvos negros desarmados, replicando achados anteriores (Correll et al., 2002; Correll et al., 2006; Payne &

Correll, 2020). Contudo, não emergiu interação entre raça do alvo e condição experimental. Esse resultado representa uma divergência potencialmente interessante entre métricas de acurácia e de velocidade de resposta. Mesmo após uma intervenção que parece alterar padrões enviesados de resposta, os tempos de reação permaneceram mais sensíveis a contingências sociais amplas do que às contingências imediatas da intervenção. Evidentemente, sob uma perspectiva analítico-comportamental, respostas mais rápidas a alvos negros podem refletir controle de estímulos moldado por exposição prolongada a comunidades verbais culturalmente consistentes, em vez de preconceito automático ou encoberto. Essas respostas relacionais historicamente reforçadas podem persistir mesmo após mudanças nas contingências que governam o comportamento manifesto. Nesse sentido, embora as intervenções possam não alterar imediatamente os padrões temporais de resposta, elas parecem, neste caso, interromper efetivamente a probabilidade de respostas enviesadas. Essa interpretação é consistente com a teoria do *behavioral momentum* (Nevin et al., 1983; Nevin & Grace, 2000; Nevin & Shahan, 2011), que postula que comportamentos historicamente reforçados exibem resistência à mudança. Mais especificamente, à medida que um operante relacional complexo passa por mudanças em função do encontro com contingências de reforçamento passadas e atuais, é esperado que o período de transição envolva a desestabilização de diferentes aspectos do operante previamente estabelecido (por exemplo, acurácia, velocidade) em ritmos distintos. Estudos futuros poderiam investigar se intervenções prolongadas ou repetidas levam a uma convergência progressiva dessas características do padrão de resposta relacional ao longo do tempo.

As análises dos índices da Teoria da Detecção de Sinais ( $c$  e  $d'$ ) revelaram diferenças de viés racial no desempenho da tarefa entre condições. Nos grupos Controle e Consistente com Estereótipos, os participantes exibiram um critério de decisão de disparo mais conservador ao responder a alvos brancos e um critério mais liberal ao responder a alvos negros. Esse padrão foi acompanhado por maior sensibilidade ( $d'$ ) para alvos brancos em comparação a alvos negros, indicando melhor discriminação entre armas e objetos não letais quando o alvo era branco. Em contraste, as intervenções de Flexibilidade e Inconsistente com Estereótipos não apresentaram diferenças em critério ou  $d'$  entre alvos negros e brancos, indicando um controle de estímulos mais simétrico por armas e objetos não letais, que não se modulou em função do sinal de raça do alvo. Importante destacar que, de acordo com nossa abordagem contextual-comportamental, tanto  $d'$  quanto  $c$  são interpretados aqui como medidas de histórias de reforçamento social, e não como métricas de traços latentes ou estados cognitivos internos (ver também Critchfield, 2002; Critchfield & Perone, 1993; Roche & Barnes-Holmes, 2003). De fato, o presente estudo chama atenção para os méritos da Teoria da Detecção de Sinais como abordagem para quantificar a discriminação de estímulos, conforme discutido em outros trabalhos (Critchfield & Perone, 1993; Lerman et al., 2010). Em vez de indexar construtos cognitivos latentes, o  $d'$  indexa o grau de controle exercido pelos objetos estímulo, enquanto o critério ( $c$ ) fornece uma medida do grau de controle exercido por pistas contextuais sobre as discriminações de estímulos.

No que diz respeito às medidas explícitas de viés racial, os resultados revelaram, de forma inesperada, um relato geral de viés racial invertido na MMRS,

bem como avaliações mais positivas de faces negras em comparação com faces brancas nas escalas Likert. A ausência de convergência entre medidas implícitas e explícitas é amplamente documentada tanto na psicologia social (Dasgupta & Greenwald, 2001; Fazio & Olson, 2003; Greenwald et al., 2009; Hofmann et al., 2005; Payne, 2001; Sawyer & Gampa, 2018) quanto na análise do comportamento (Cartwright et al., 2016; Passarelli, Mizael et al., 2024; Roche et al., 2024; Roche et al., 2025). Uma fonte evidente dessa discrepância é que medidas de autorrelato geralmente não impõem restrições de tempo e são suscetíveis a contingências atuais, incluindo vieses de desejabilidade social (Krumpal, 2013). De fato, as características sociodemográficas da amostra (isto é, indivíduos mais jovens e com maior nível educacional) podem ter contribuído para um viés de apresentação mais forte na MMRS (Branton & Jones, 2005; Greeley & Sheatsley, 2017; Oliver & Mendelberg, 2000).

Independentemente da fonte da discrepância, esse resultado fragiliza a confiança em autorrelatos de viés racial, dado o número recorrente de observações de viés racial na PODT na literatura, bem como no presente estudo com os mesmos participantes. Embora o viés exibido na PODT de forma alguma indique que um indivíduo responderia da mesma maneira em uma situação real (ver Cesario, 2022), o desempenho na PODT representa, ainda assim, uma forma de viés em si, ainda que em um ambiente artificial, como um jogo. Embora seja comum na literatura sociocognitiva reconhecer que o conceito relativamente difuso de “atitude” é definido em termos de componentes comportamentais, cognitivos e afetivos (Rosenberg & Hovland, 1960), que nem sempre covariam perfeitamente, não conhecemos modelos nos quais diferentes componentes do conceito de

atitude se oponham diametralmente entre si. Claramente, compreender a relação entre diferentes aspectos do viés relacional (isto é, explícito versus implícito) é uma tarefa complexa (ver Corneille & Gawronski, 2024), agravada pela definição imprecisa do próprio conceito de atitude, especialmente do ponto de vista da análise do comportamento (ver Watters et al., 2023).

Embora não tenha havido correspondência sistemática entre autorrelatos explícitos e a medida implícita de viés racial, análises correlacionais exploratórias revelaram padrões potencialmente significativos. Dado que múltiplas correlações foram realizadas sem correção para o erro do tipo I, o pequeno número de associações estatisticamente significativas deve ser interpretado com cautela. Ainda assim, as análises indicaram a possibilidade de algumas relações entre medidas explícitas e implícitas em aspectos específicos. Por exemplo, tanto nas condições consistente quanto inconsistente com estereótipos, escores mais baixos na MMRS (indicativos de atitudes pró-negros) covariaram com maior acurácia em não atirar em alvos negros desarmados. Na condição inconsistente com estereótipos, escores mais baixos na MMRS também estiveram associados a maior discriminabilidade ( $d'$ ) para alvos negros, indicando controle de estímulos mais preciso à medida que a adesão explícita a atitudes racistas diminuía. Assim, embora não tenha sido observada covariância global entre medidas explícitas e implícitas, que frequentemente é modesta mesmo quando encontrada (por exemplo, Greenwald et al., 2009; Hofmann et al., 2010), os resultados sugerem algumas relações pontuais entre essas medidas em determinados aspectos.

Importante destacar que as análises correlacionais também revelaram que o viés racial na PODT não se limitou às respostas de disparo contra alvos negros,

mas também envolveu padrões de respostas de disparo e não disparo em relação a alvos brancos. Mais especificamente, na intervenção de flexibilidade, maiores reduções nos escores de viés RFD ao longo das administrações completas do FAST (isto é, inclinação do RFD) correlacionaram-se negativamente com os tempos de resposta de não disparo para alvos brancos (ou seja, maior redução de viés no FAST previu respostas menos enviesadas positivamente em relação a alvos brancos). Além disso, no grupo controle, avaliações mais positivas de faces negras covariaram com tempos de reação mais rápidos para disparos corretos em alvos brancos armados, um padrão consistente com autorrelatos de preferência por faces negras. Esses achados sugerem que existe uma relação complexa e multifacetada entre vieses pró-negros, antinegros, pró-brancos e antibranco. Consequentemente, intervenções voltadas à redução do viés racial devem adotar uma abordagem holística que contemple todos esses aspectos (Lai et al., 2016; Passarelli, Roche et al., 2024).

Esforços futuros de intervenção podem se beneficiar da explicitação de um importante paralelo entre a conceitualização atual do FAST e processos terapêuticos centrais descritos por analistas do comportamento e terapeutas comportamentais. Especificamente, o FAST tem se mostrado uma ferramenta eficaz para capturar a dinâmica do *momentum* do responder relacional (Cartwright et al., 2016; Cummins et al., 2018; Cummins & Roche, 2020; Passarelli, Mizael et al., 2024) e a resistência à mudança de relações verbais. Esses conceitos estão intimamente relacionados à flexibilidade relacional, considerada um processo psicológico-chave para a mudança terapêutica. A inflexibilidade psicológica pode ser entendida como a tendência a evitar ou resistir à modificação de crenças e

atitudes, mesmo diante de evidências contraditórias (Levin et al., 2016; Florez et al., 2019), sendo por vezes caracterizada como “fusão cognitiva”, na qual pensamentos e crenças são tratados como verdades literais, impedindo a adaptação eficaz a novos contextos ou informações (Macri & Rogge, 2024). Por exemplo, indivíduos podem continuar relacionando pessoas negras a termos negativos apesar da exposição a inúmeros exemplos positivos e a estratégias de reforçamento destinadas a contrariar essas associações.

Por outro lado, a flexibilidade psicológica, definida como a capacidade de se adaptar às demandas situacionais e perseguir objetivos significativos mesmo em situações desafiadoras, é crucial para promover mudanças em atitudes raciais (Hayes, 2004; Doorley et al., 2020). A técnica de “desfusão cognitiva”, que envolve distanciar-se dos próprios pensamentos e responder com flexibilidade relacional sensível ao contexto, pode ser uma abordagem eficaz para modificar preconceitos (Farmer et al., 2023; Hayes et al., 2004). Estudos baseados na Terapia de Aceitação e Compromisso (ACT) têm mostrado resultados promissores ao auxiliar indivíduos a transformar padrões relacionais negativos, ampliando as funções dos estímulos e aumentando o repertório de respostas em situações sociais (Farmer et al., 2023; Levin et al., 2016; Lillis & Hayes, 2007; Masuda et al., 2004).

O presente estudo demonstra que o FAST pode ser uma ferramenta valiosa para promover flexibilidade relacional e, conseqüentemente, favorecer a “desfusão cognitiva” por meio de um procedimento automatizado e digital. Um estudo prévio sugeriu, inclusive, que a ferramenta pode ser útil no tratamento de cognições rígidas associadas à dor crônica (Jacobsen et al., 2020). No presente contexto, o aumento da flexibilidade relacional mostrou-se parcialmente eficaz na mitigação

do viés de disparo na PODT. Curiosamente, a exposição dos participantes a intervenções destinadas a reforçar respostas consistentes ou inconsistentes com estereótipos também pareceu direcionar os vieses de resposta na direção correspondente. Estudos anteriores em psicologia social experimental igualmente demonstraram que a exposição repetida a contra exemplares é eficaz na redução de vieses raciais implícitos no curto prazo (Dixon et al., 2005; Forscher et al., 2019; Lai et al., 2014, 2016).

Também é digno de nota que os achados atuais sugerem que, para promover a “desfusão cognitiva”, pode ser tão eficaz expor repetidamente indivíduos a treino relacional inconsistente com estereótipos quanto expô-los a treino explícito de flexibilidade. A aparente eficácia ampla de ambos os métodos é interessante por duas razões. Primeiro, a flexibilidade relacional é considerada um processo comportamental central subjacente à mudança terapêutica e, no contexto clínico, costuma ser priorizada em relação a tentativas de modificar diretamente o comportamento atual. Assim, é surpreendente que o simples confronto de estereótipos pareça operar de forma comparável a uma intervenção de flexibilidade na redução do preconceito, embora essa equivalência não tenha sido testada estatisticamente.

Uma interpretação da eficácia da intervenção inconsistente com estereótipos reside no fato de que a maioria dos participantes provavelmente iniciou o estudo com uma história pré-experimental de exposição a viés antinegro, o que pode ter tornado a condição inconsistente funcionalmente equivalente à condição de flexibilidade. Em termos simples, o contato com treino relacional contraestereotípico, dada uma longa história de responder relacional enviesado

racialmente, pode funcionar como um exercício de defusão, no qual o indivíduo é desafiado a considerar perspectivas alternativas (Devine et al., 2012; Forscher et al., 2019; Lai et al., 2014; Kawakami et al., 2007). Embora a mudança comportamental significativo costume requerer mudanças contextuais (Hayes et al., 2006; Hayes et al., 2012), é possível que o treino inconsistente com estereótipos represente, por si só, uma mudança de contexto suficiente e eficaz, tornando desnecessárias intervenções explícitas adicionais de flexibilidade e defusão.

Em segundo lugar, é notável que a intervenção de flexibilidade utilizada tenha produzido efeitos apesar de envolver, em igual proporção, o reforçamento e a contradição direta de estereótipos raciais. O fato de essa intervenção ter sido profilaticamente eficaz sugere fortemente que a flexibilidade relacional é um processo potencialmente poderoso, uma vez que reduziu o viés na PODT na ausência de qualquer característica da intervenção voltada diretamente à eliminação do viés racial. Em termos simples, metade das tentativas na intervenção de flexibilidade envolveu o reforçamento de estereótipos raciais e, ainda assim, o responder estereotipado foi reduzido com sucesso.

Apesar dos achados promissores, o presente estudo apresenta limitações metodológicas e conceituais. Primeiramente, a ausência de uma linha de base na PODT pode ter comprometido a avaliação de mudanças individuais no preconceito ao longo do tempo. Contudo, a decisão de não registrar uma linha de base foi tomada deliberadamente para evitar comprometer a integridade da medida subsequente em função de efeitos de prática. De modo semelhante, o estudo não incluiu uma avaliação de seguimento de longo prazo, o que limita a avaliação da

durabilidade das mudanças comportamentais (Lai et al., 2014, 2016). Avaliações de seguimento seriam menos suscetíveis a efeitos de prática do que a inclusão de uma linha de base e, portanto, constituem a opção mais adequada para pesquisas futuras, por possibilitarem uma compreensão mais abrangente da eficácia e persistência dos efeitos da intervenção (Forscher et al., 2019; Lai et al., 2016; Passarelli, Roche et al., 2024).

Uma segunda limitação metodológica refere-se ao sistema de pontuação gamificado e relativamente arbitrário empregado na PODT, concebido como um esquema de incentivo para o engajamento dos participantes. Embora os pontos atribuídos ou subtraídos não tenham sido utilizados como variáveis dependentes, a penalização de 40 pontos por não atirar em um alvo armado pode ter aumentado a vigilância dos participantes em relação a alvos armados, possivelmente explicando a ausência de viés racial significativo no grupo controle. Embora esse procedimento tenha sido mantido para replicar fielmente o método de Correll et al. (2002), estudos futuros poderiam considerar a remoção ou modificação desse esquema de incentivo. Tal modificação poderia ajudar a compreender o potencial da PODT para captar discriminação racial para além do contexto específico do dilema policial, permitindo observar respostas enviesadas de forma mais natural, sem a influência de um sistema de pontuação arbitrário.

Por fim, questionamentos têm sido levantados acerca da validade ecológica de tarefas como a PODT e de sua capacidade de prever o comportamento individual em cenários complexos do mundo real (Cesario, 2022). No entanto, a principal utilidade desses paradigmas laboratoriais não reside na predição direta de outros comportamentos, mas em sua capacidade de isolar e manipular

experimentalmente componentes centrais do comportamento discriminatórios este caso, a tomada de decisão rápida sob pressão temporal. O robusto efeito de viés capturado pela PODT em laboratório é consistente com disparidades bem documentadas na percepção de ameaça durante abordagens policiais (Carvalho et al., 2022), com o viés racial em decisões de disparo (Correll et al., 2002, 2007, 2014) e com vieses raciais avaliativos implícitos (Ahadinezhad et al., 2022; Passarelli, Mizael et al., 2024; Roche et al., 2025). Isso sustenta a ideia de que a PODT representa um análogo persuasivo de pontos de escolha de resposta socialmente situados, útil para o estudo dos mecanismos comportamentais que mantêm vieses sociais persistentes.

Os achados do presente estudo oferecem um caminho proativo que conecta a pesquisa comportamental básica a intervenções de relevância social prática. Essa abordagem fundamenta-se no princípio comportamental bem estabelecido de que o responder relacional derivado pode ser modificado por meio de treino baseado em contingências (Mizael et al., 2016, 2022; Passarelli et al., 2024; Romani et al., 2024; Suarez et al., 2024). Ao incorporar um componente ativo de treino com reforçamento direto, o FAST ultrapassa o escopo meramente avaliativo das tarefas tradicionais de viés implícito e alinha-se aos objetivos interventivos da Terapia de Aceitação e Compromisso (ACT), que visa cultivar flexibilidade comportamental diante de padrões rígidos de resposta relacionados ao viés (Florez et al., 2019; Lillis & Hayes, 2007; Romani et al., 2024). Consequentemente, este trabalho oferece uma metodologia funcional e escalável para o estabelecimento de repertórios relacionais flexíveis e não preconceituosos, com potencial direto para o desenvolvimento de programas baseados em

evidências destinados à mitigação de vieses em contextos de alto impacto social, como aplicação da lei, educação, tomada de decisão judicial e avaliação clínica.

-----

## **Comentário Final do Capítulo**

### **O Espelho e o Martelo: Refletindo Padrões Consistentes e Quebrando-os com a Variação Inconsistente**

Esse estudo demonstrou como diferentes arranjos de contingências selecionam, mantêm ou enfraquecem padrões de responder relacional com viés racial antinegro. Nesse sentido, o valor principal do experimento está menos no rótulo da intervenção e mais na comparação funcional entre condições que reforçam, desafiam ou alternam relações simbólicas culturalmente predominantes.

Um achado particularmente robusto foi a eficácia do treino inconsistente com estereótipos. A exposição repetida a relações contra estereotípicas mostrou-se suficiente para reduzir vieses raciais de decisão na tarefa de tomada de decisão envolvendo atirar ou não em alvos armados e desarmados sob pressão de tempo. Do ponto de vista comportamental, o contato sistemático com contingências inconsistentes pode funcionar como uma forma de ‘desmantelar’ ou confrontar o repertório dominante, promovendo reorganização relacional sem necessidade de instruções metacognitivas ou apelos normativos.

Um ponto conceitualmente relevante deste estudo emerge da comparação entre o treino baseado em flexibilidade relacional e o treino inconsistente com estereótipos. Havia a expectativa de que a alternância sistemática entre padrões consistentes e inconsistentes promovesse maior mitigação do viés racial, ao ampliar a variabilidade do responder. No entanto, os dados mostraram que o treino exclusivamente inconsistente produziu efeitos similares e, em algumas medidas, até mais robustos, do que a condição de flexibilidade.

Esses resultados sugerem que essas condições não operam como estratégias concorrentes, mas como variações funcionais de um mesmo processo. O treino

inconsistente, ao reforçar respostas específicas na presença de estímulos raciais, estabelece novas relações funcionais entre estímulos, respostas e consequências. Conforme argumentado por Sidman (2000), quando respostas são diferencialmente reforçadas em presença de determinados estímulos, elas podem passar a integrar as classes de equivalência formadas.

Sob essa perspectiva, o treino inconsistente no FAST não atua apenas como um procedimento “avaliativo” ou de oposição ao estereótipo, mas como um arranjo capaz de produzir classes relacionais funcionalmente equivalentes às estabelecidas em procedimentos clássicos de equivalência de estímulos. A diferença reside menos na lógica comportamental envolvida e mais na forma procedimental pela qual as relações são estabelecidas.

Essa interpretação ajuda a compreender por que o treino inconsistente foi eficiente para promover reorganização relacional significativa, mesmo sem alternância explícita entre padrões. Assim, os dados indicam que tanto o treino de flexibilidade quanto o treino inconsistente podem produzir mitigação do viés, desde que estabeleçam contingências capazes de integrar estímulos raciais, respostas e funções avaliativas em novas organizações relacionais. O que se mostra decisivo, portanto, não é a alternância em si, mas o estabelecimento de relações que passem a competir, funcionalmente, com redes simbólicas culturalmente dominantes.

Em contraste, a condição consistente com estereótipos operou como uma demonstração experimental dos efeitos danosos da repetição cultural de contingências predominante racistas. Nessa condição, os participantes exibiram padrões claros de viés antinegro na acurácia, no critério de decisão e na sensibilidade discriminativa. Longe de representar um “grupo controle ativo”, essa condição ilustra como a simples prática

reiterada de relações simbólicas socialmente reforçadas é suficiente para fortalecer padrões enviesados de tomada de decisão. Em termos funcionais, trata-se de um análogo experimental das práticas culturais que, ao longo do tempo, selecionam e mantêm respostas discriminatórias mesmo em indivíduos que explicitamente rejeitam o racismo.

A tarefa do dilema do policial (PODT) ocupa um papel central nessa análise justamente por se tratar de uma medida que aproxima o laboratório de contextos de decisão socialmente relevantes. Embora o cenário específico do porte de armas não seja típico do contexto irlandês, a tarefa modela um ponto crítico de escolha sob pressão temporal, ambiguidade e consequências diferenciais, características comuns a múltiplos contextos naturais de tomada de decisão. Seu caráter lúdico não diminui sua relevância analítica; ao contrário, permite observar como relações historicamente estabelecidas entram em controle quando o tempo para deliberação é restringido.

As métricas derivadas da Teoria da Detecção de Sinais, acurácia, falsos alarmes, critério (c) e sensibilidade ( $d'$ ), mostram-se particularmente úteis quando interpretadas à luz da análise do comportamento. Longe de indexarem “viés interno” ou “processos inconscientes”, essas medidas podem ser compreendidas como produtos diretos de histórias diferenciais de reforçamento e controle contextual. O critério reflete a tendência global de responder sob determinadas contingências, enquanto o  $d'$  expressa o grau de controle exercido pelos estímulos relevantes da tarefa, independentemente da raça. A abertura da análise do comportamento a esse tipo de métrica amplia o diálogo metodológico sem comprometer o compromisso funcional da interpretação.

Um aspecto recorrente e teoricamente relevante foi a dissociação entre medidas explícitas e implícitas. Os participantes, de modo geral, contornaram com sucesso medidas explícitas de racismo moderno, exibindo padrões compatíveis com normas sociais

contemporâneas de rejeição ao preconceito. No entanto, esse controle verbal não se estendeu de forma consistente às medidas sob pressão temporal. Esse descompasso não deve ser interpretado como contradição, mas como evidência de que diferentes contingências controlam diferentes classes de resposta. Uma questão que permanece em aberto é se os padrões observados nas medidas explícitas se alterariam em condições de anonimato garantido, nas quais o controle por audiência e deseabilidade social seria reduzido. Estudos futuros deveriam manipular sistematicamente essa variável para separar o controle exercido por contingências sociais imediatas do controle exercido pela história relacional dos participantes. O responder explícito parece mais sensível a contingências sociais imediatas, enquanto o responder sob pressão temporal revela a persistência de relações simbólicas historicamente reforçadas.

Em conjunto, os dados reforçam uma conclusão central: a modificação do viés racial não depende apenas da substituição de conteúdos avaliativos, mas da reorganização funcional do responder relacional. O FAST mostrou-se uma ferramenta potente para tornar visível tanto a rigidez quanto a flexibilidade desse responder, funcionando simultaneamente como medida e como arranjo de treino. Mais do que oferecer uma solução, o estudo explicita os limites e as possibilidades do controle experimental sobre fenômenos moldados culturalmente, reafirmando uma postura científica cautelosa e analítica: compreender o comportamento implica identificar as contingências que o organizam, inclusive quando essas contingências produzem efeitos socialmente indesejáveis.

\*\*\*\*\*

### **Considerações finais**

Esta tese se encerra com a expectativa de contribuir para a consolidação de agendas de pesquisa mais sistemáticas sobre preconceito racial no âmbito da análise do comportamento. Se, por um lado, o racismo pode ser compreendido como uma “caixa de Pandora”, espera-se que a ciência funcione aqui como instrumento de análise funcional e transformação, e não como mais um mecanismo de exclusão ou simplificação indevida do fenômeno.

As considerações finais organizam um conjunto de sínteses analítico-funcionais que emergiram do contato reiterado com fluxo analítico, teórico e empírico formulados na tese. Esses pontos não devem ser lidos como prescrições normativas, mas como implicações funcionais decorrentes da articulação entre teoria, procedimentos, medidas e resultados experimentais.

#### **Síntese 1 – Necessidade da sistematização de uma agenda de pesquisa sobre racismo e preconceito racial na Análise do Comportamento**

Uma constatação importante desta tese diz respeito às próprias práticas da análise do comportamento. Apesar da existência de contribuições relevantes e conceitualmente sofisticadas, o campo possui uma agenda de pesquisa incipiente, de modo que os estudos têm sido conduzidos de forma fragmentada, com pouco diálogo funcional entre procedimentos, medidas e níveis de análise.

Essa fragmentação dificulta a acumulação de conhecimento e favorece a repetição de equívocos conceituais e metodológicos. A construção lógica dessa tese parece indicar que avanços consistentes nesse campo exigem menos estudos isolados e mais programas de pesquisa integrados, organizados explicitamente em torno de processos comportamentais bem definidos.

## **Síntese 2 - Formulações teóricas não devem anteceder as observações experimentais**

Uma diretriz metodológica que atravessa toda a tese é a recusa de formulações teóricas que antecedam ou constringam as observações experimentais. Em diversos momentos, os resultados contrariaram expectativas iniciais, exigindo ajustes conceituais e reformulações interpretativas.

Esse movimento não representa fragilidade teórica, mas uma escolha analítica coerente com uma tradição da ciência do comportamento que permite que os dados orientem, gradualmente, a organização conceitual, sem a imposição prévia de modelos explicativos. À maneira defendida por Sidman (1994), foi o contato continuado com os dados, e não compromissos prévios com estruturas teóricas específicas, que orientou a construção progressiva das interpretações. Em contextos como o presente, essa estratégia reduz a probabilidade de que os resultados sejam subsumidos a narrativas explicativas pré-estabelecidas e favorece a identificação de regularidades funcionais passíveis de acumulação empírica.

## **Síntese 3 - Racismo e preconceito racial não são conceitos intercambiáveis**

Uma síntese conceitual importante da tese diz respeito à distinção funcional entre racismo e preconceito racial. Embora relacionados, esses fenômenos operam em níveis distintos e não podem ser tratados como sinônimos.

O preconceito racial refere-se a repertórios comportamentais individuais, passíveis de investigação experimental em termos de controle de estímulos, história de reforçamento e flexibilidade relacional. O racismo, por sua vez, envolve práticas coletivas, nas quais contingências que operam em sistemas, instituições e práticas culturais organizam e mantêm esses repertórios, produzindo e sustentando desigualdades estruturais.

A Análise do Comportamento oferece ferramentas conceituais e metodológicas para investigar ambos os níveis, desde que a distinção funcional entre eles seja preservada. O preconceito racial pode ser analisado no nível do indivíduo. O racismo exige a análise das contingências culturais, políticas e institucionais que estruturam o ambiente social. Confundir esses níveis implica reduzir fenômenos coletivos a repertórios individuais ou, inversamente, tratar padrões individuais como explicações suficientes para processos estruturais.

Assim, a distinção entre racismo e preconceito racial não fragmenta o fenômeno, mas permite respeitar os diferentes níveis de análise envolvidos, ampliando a precisão conceitual e evitando erros funcionais na interpretação e na pesquisa.

#### **Síntese 4 – Preconceito racial como responder relacional sob múltiplos controles**

O preconceito racial não pode ser adequadamente compreendido apenas como avaliação negativa dirigida a grupos raciais. Tratá-lo exclusivamente como avaliação obscurece dimensões centrais do fenômeno, como relações de hierarquia, oposição, comparação, diferença, preferência e discriminação.

Mais do que valências atribuídas a estímulos sociais isolados, que constitui apenas mais uma dimensão desse fenômeno. o preconceito racial constitui um padrão de responder relacional, historicamente construído, mantido por contingências culturais e sensível às propriedades contextuais que organizam as relações entre estímulos raciais.

Nessa perspectiva, o preconceito racial não configura um repertório unitário. Diferentes medidas capturam diferentes formas de controle de estímulos, que podem coexistir de maneira não isomórfica no comportamento de um mesmo indivíduo. Respostas com topografia aparentemente positiva (por exemplo, relacionar pessoas negras à força

física) podem operar sob controles igualmente danosos e discriminatórios, sustentando padrões de exclusão diferencial.

Assim, diferentes medidas revelam diferentes controles comportamentais em operação, indicando que a identificação do preconceito racial exige uma análise funcional que considere a multiplicidade de relações simbólicas e contextos nos quais essas relações são evocadas e mantidas.

### **Síntese 5 – As medidas utilizadas mostraram estabilidade sob repetição espaçada**

As evidências produzidas nesta tese indicam ausência sistemática de efeitos de teste–reteste sob reaplicação espaçada das medidas de preconceito racial. Essa estabilidade foi observada tanto em medidas explícitas quanto no FAST, uma medida implícita baseada em tempo de resposta.

Esse achado contraria a suposição de que medidas implícitas não poderiam ser reaplicadas sem produzir efeitos de prática. Os dados sugerem que, quando reaplicadas com espaçamento temporal adequado, as medidas permanecem funcionalmente estáveis. A ausência de efeito de teste–reteste não deve ser confundida com efeito de prática, que requer repetição contínua e produção de curvas de aquisição.

### **Síntese 6 – A idade da amostra mostrou-se uma variável funcionalmente relevante**

Nos estudos com crianças, participantes mais jovens apresentaram maior correspondência entre diferentes medidas de preconceito racial. Esse padrão sugere menor influência de filtros sociais e de repertórios de desajustabilidade social, frequentemente observados em faixas etárias mais avançadas. Em crianças mais novas, as respostas tenderam a refletir de forma mais direta as relações funcionais em vigor, tornando mais visível o controle exercido por estímulos raciais nas diferentes tarefas. A idade modulou não apenas a intensidade do viés, mas também a convergência entre medidas.

Do ponto de vista interventivo, esses resultados indicam que contextos nos quais o controle social sobre “o que deve ser dito” ainda não está plenamente estabelecido constituem oportunidades privilegiadas para a reorganização de relações simbólicas. Intervenções que operam diretamente sobre contingências relacionais, antes da consolidação de repertórios de autocontrole verbal e desejabilidade social, tendem a produzir efeitos mais consistentes e generalizáveis.

Assim, em vez de focalizar a correção verbal ou a conscientização explícita, práticas interventivas mais promissoras são aquelas que alteram sistematicamente as contingências que organizam o responder relacional, especialmente em fases do desenvolvimento em que o comportamento ainda está menos mediado por regras sociais abstratas. Esses achados sugerem que a efetividade da intervenção não depende apenas do procedimento empregado, mas também do momento do desenvolvimento em que ele é introduzido.

### **Síntese 7 – A formação de classes de equivalência é condição necessária para mudanças, mas não garante generalização**

Ao longo da tese, observou-se de forma sistemática que, quando classes de equivalência foram formadas, houve eliminação das frequências de respostas que relacionavam faces negras a símbolos negativos, bem como redução consistente das diferenças de valência atribuídas a faces negras e brancas. Esses efeitos foram replicados em diferentes estudos, tarefas e medidas, indicando mudanças robustas no responder relacional e avaliativo.

Na ausência da formação de classes de equivalência, nenhuma dessas mudanças foi observada. Esse padrão indica que a formação de classes constitui um pré-requisito funcional para a mitigação do viés antinegro, tanto no responder relacional quanto nas avaliações de valência.

Entretanto, a formação dessas classes não foi suficiente para assegurar generalização ampla dos efeitos para todas as medidas e para todos os participantes. Identificou-se um grupo que apresentou mudanças intermediárias, caracterizadas por generalização parcial dos efeitos do treino: nesses casos, ocorreram alterações em algumas tarefas ou medidas, enquanto a raça permaneceu exercendo controle sobre o responder em contextos específicos.

Assim, embora as relações de equivalência tenham sido estabelecidas, o controle relacional racial não foi completamente enfraquecido nesses participantes. Esse padrão sugere que a generalização das mudanças depende não apenas da formação de classes, mas também da história prévia de controle racial e da força funcional desse controle, que pode persistir mesmo após reorganização das relações simbólicas.

#### **Síntese 8 – Resistência à mudança emerge cedo**

Mesmo em crianças pequenas, com seis anos, observou-se resistência à reorganização relacional. Isso sugere que práticas culturais amplas organizam o comportamento racial desde cedo.

#### **Síntese 9 – Flexibilidade relacional envolve eliminar a raça como controle do responder**

Neste estudo, a flexibilidade do responder relacional foi caracterizada principalmente pela redução do controle do fenótipo dos estímulos sobre o comportamento, e não pela formação de novos padrões relacionais estáveis. Flexibilizar o responder significou, empiricamente, deixar de utilizar características fenotípicas como dimensão discriminativa dominante, especialmente quando organizadas de forma dicotômica ou de oposição.

Quando esse controle foi enfraquecido, os participantes passaram a responder de maneira menos sistemática às categorias raciais, indicando perda da função discriminativa da raça no controle do responder relacional. Essa mudança não envolveu a substituição por novas relações simbólicas claramente definidas, mas a diminuição do peso funcional da raça na organização do comportamento.

Esse padrão foi exemplificado no Teste da Boneca, no qual alguns participantes, sem qualquer treino ou instrução prévia, passaram a realizar escolhas sociais não excludentes, selecionando todas as figuras disponíveis, sem rejeição sistemática de nenhuma delas. Esses participantes também apresentaram mudanças convergentes em outras medidas, sugerindo que a flexibilização observada não se restringiu a uma tarefa específica, mas refletiu uma alteração mais ampla no controle relacional do comportamento.

### **Síntese 10 – Validade social não é acessório, é a construção da relação da pesquisadora com todos os membros do ambiente da pesquisa**

Mudanças experimentais nem sempre se traduziram em efeitos observáveis no cotidiano. Estudos longitudinais, com inserção institucional e articulação com equipes escolares, apresentaram maior validade social.

### **Síntese 11 – FAST pode ser usado como medida, mas também como ferramenta de treino**

Os resultados desta tese indicam que o *Function Acquisition Speed Test* (FAST) não deve ser compreendido apenas como um instrumento de mensuração do viés racial, mas também como um procedimento capaz de produzir mudanças no controle relacional quando utilizado de forma sistemática. Ao operar sob restrição temporal e exigir respostas rápidas, o FAST indexa a força de relações simbólicas previamente estabelecidas; ao mesmo tempo,

a exposição reiterada a contingências específicas pode alterar essas relações, especialmente quando o procedimento é configurado de modo a favorecer relações contra estereotípicas ou flexíveis.

Esse duplo estatuto aproxima o FAST de outros procedimentos relacionais, como o *Implicit Relational Assessment Procedure* (IRAP), que também tem sido utilizado tanto para avaliar a força do responder relacional quanto como ferramenta de treino. Estudos utilizando o *Implicit Relational Assessment Procedure* (IRAP) demonstraram que manipulações sistemáticas nas contingências da tarefa podem não apenas mensurar, mas também alterar a força e a direção de relações derivadas, enfraquecendo padrões dominantes e fortalecendo relações alternativas (Harte et al., 2021; Barnes-Holmes & Harte, 2022). Esses achados indicam que medidas implícitas não são funcionalmente neutras, podendo operar como contextos de treino relacional. De modo similar, os resultados desta tese sugerem que o *Function Acquisition Speed Test* (FAST) pode exercer dupla função: avaliar o viés relacional e, sob determinadas configurações contingenciais, participar ativamente de sua modificação.

Do ponto de vista aplicado, essa característica amplia o potencial do FAST, permitindo concebê-lo como parte de intervenções baseadas na reorganização de contingências relacionais, e não apenas como um instrumento de medida. Ao mesmo tempo, esses achados reforçam a necessidade de cautela metodológica: quando utilizado como medida, é essencial considerar os efeitos da exposição repetida; quando utilizado como treino, é preciso planejar deliberadamente os parâmetros que maximizam a probabilidade de mudança. Em ambos os casos, o FAST se insere como uma ferramenta versátil dentro de uma abordagem analítico-comportamental do viés racial.

## **Síntese 12 – Reorganização relacional é funcionalmente distinta da formação de classes**

Os estudos desta tese não investigaram formação de classes de equivalência em sentido estrito, mas reorganização relacional, um processo funcionalmente mais exigente. Nos procedimentos clássicos de equivalência, estímulos arbitrários e sem história prévia são relacionados experimentalmente, e classes emergem sem competição com repertórios pré-existentes. Nos estudos aqui apresentados, os estímulos raciais já integravam classes funcionais culturalmente estabelecidas antes da intervenção: faces negras participavam de redes simbólicas organizadas por coordenação com atributos negativos e por oposição a atributos positivos.

O procedimento experimental não criou classes novas a partir do zero — ele tentou mover estímulos de uma classe para outra, o que exige simultaneamente o estabelecimento de novas relações e o enfraquecimento das relações anteriores. Essa distinção tem implicações diretas para a interpretação dos resultados: a resistência à mudança observada, a generalização parcial e a ambivalência funcional, como o mesmo estímulo participando de classes positivas e negativas — não são falhas do procedimento, mas expressões esperadas de um processo de reorganização que opera sobre redes simbólicas historicamente consolidadas. Compreender o preconceito racial como problema de reorganização relacional, e não de formação, reposiciona as perguntas centrais do campo: não se trata de saber se equivalência pode ser estabelecida, mas sob quais condições relações previamente estabelecidas podem ser suficientemente enfraquecidas para que novas organizações funcionais se tornem estáveis.

### **Síntese 13 – A exposição às contingências que reforçam viés racial antinegro aprofunda o viés**

A exposição reiterada a contingências consistentes com estereótipos raciais negativos fortaleceu padrões de viés antinegro na tomada de decisão em uma tarefa baseada no Dilema do Policial. Nessas condições, os participantes apresentaram maior tendência a decisões enviesadas sob restrição temporal, indicando que relações simbólicas estereotípicas exerceram controle direto sobre respostas rápidas de julgamento e ação.

Esse efeito pode ser compreendido por analogia a práticas midiáticas recorrentes. A exposição repetida a noticiários que associam pessoas negras à criminalidade ou a narrativas ficcionais que posicionam pessoas negras em papéis subalternizados funciona, do ponto de vista comportamental, como um treino contínuo de relações simbólicas estereotípicas. Essas contingências não instruem explicitamente o preconceito, mas organizam contextos nos quais determinadas respostas são sistematicamente selecionadas.

De modo funcionalmente análogo, a condição experimental consistente estereotípica simulou um ambiente no qual decisões rápidas foram continuamente orientadas por relações raciais previamente estabelecidas, resultando no aprofundamento do viés em comparação ao grupo controle. Essa analogia sugere que intervenções eficazes precisam competir com esse tipo de treino ambiental crônico, não apenas por meio de mensagens contrárias, mas pela reorganização ativa das contingências que sustentam o controle relacional racial no cotidiano.

### **Síntese 14 – Avançar exige precisão conceitual, controle experimental e humildade empírica**

O estudo do preconceito racial exige rigor conceitual, método experimental refinado e disposição para rever pressupostos à luz dos dados. A contribuição da análise do

comportamento está em identificar processos, limites e condições de mudança, não em prometer soluções universais.

### **Síntese 15 –Intervir é manipular contingências, não corrigir pessoas**

A síntese final desta tese é que intervenções eficazes sobre o preconceito racial operam por meio da manipulação de contingências, e não pela tentativa de corrigir pessoas, atitudes ou crenças. As mudanças observadas ao longo dos estudos emergiram quando contextos experimentais passaram a selecionar novos padrões de responder relacional, e não quando os participantes foram expostos a apelos morais, instruções abstratas ou tentativas de conscientização verbal.

Embora esta tese não ofereça um conjunto exaustivo de respostas sobre como intervir sobre o viés racial, os dados indicam algumas condições sob as quais mudanças podem ocorrer. Em particular, procedimentos que reorganizam relações simbólicas previamente estabelecidas, que operam sob restrição temporal, ou que promovem flexibilidade relacional mostraram-se capazes de enfraquecer o controle exercido por categorias raciais sobre o comportamento, ainda que de forma gradual e heterogênea.

Esses resultados não implicam que tais procedimentos constituam as únicas ou necessariamente as melhores formas de intervenção, mas sugerem que abordagens baseadas na seleção de respostas por contingências específicas podem produzir efeitos mensuráveis onde estratégias puramente discursivas tendem a falhar. Assim, a principal contribuição desta tese não é prescrever o que deve ser feito, mas indicar algumas maneiras pelas quais a intervenção comportamental sobre o preconceito racial pode ser empiricamente explorada e refinada.

## Referências

- About, F. E. (1993). The developmental psychology of racial prejudice. *Transcultural Psychiatric Research Review*, 30(3), 229-242.  
<https://doi.org/10.1177/136346159303000303>
- About, F.E. (1988). *Children and prejudice*. Blackwell.
- About, F.E., Mendelson, M., & Purdy, K. (2003). Cross-race peer relations and friendship quality. *International Journal of Behavioral Development*, 27(2), 165–173.  
<https://doi.org/10.1080/01650250244000164>
- Adams, B. J., Fields, L., & Verhave, T. (1993). Effects of test order on intersubject variability during equivalence class formation. *The Psychological Record*, 43(1), 133–152. <https://psycnet.apa.org/record/1993-28235-001>
- Ahadinezhad, B., Khosravizadeh, O., Maleki, A., & Hashtrودي, A. (2022). Implicit racial bias among medical graduates and students by an IAT measure: A systematic review and meta-analysis. *Irish Journal of Medical Science*, 191(4), 1941-1949.  
<https://doi.org/10.1007/s11845-021-02756-3>
- Alexander, T., & Carter, M. M. (2022). Internalized racism and gendered colorism among African Americans: A study of intragroup bias, perceived discrimination, and psychological well-being. *Journal of African American Studies*, 26(2), 248-265.  
<https://doi.org/10.1007/s12111-022-09586-2>
- Almeida, S. L. D. (2019). *O que é racismo estrutural?* Belo Horizonte (MG)
- Amd, M., Machado, A., De Oliveira, M. A., Passarelli, D. A., & De Rose, J. C. (2019). Effects of nodal distance on conditioned stimulus valences across time. *Frontiers in Psychology*, 10, 742. <https://doi.org/10.3389/fpsyg.2019.00742>

- Amd, M., de Oliveira, M. A., Passarelli, D. A., Balog, L. C., & de Rose, J. C. (2018). Effects of orientation and differential reinforcement II: Transitivity and transfer across five-member sets. *Behavioural Processes, 150*, 8-16.  
<https://doi.org/10.1016/j.beproc.2018.02.012>
- Amd, M., & Passarelli, D. A. (2020). Dissociating preferences from evaluations following subliminal conditioning. *Acta Psychologica, 204*, 103023.  
<https://doi.org/10.1016/j.actpsy.2020.103023>
- Araújo, E. M. de, Xavier, K. A. da S., Souza, L. B. de, & Vichi, C. (2022). Racismo internalizado: Uma perspectiva analítico-comportamental. *Perspectivas em Análise do Comportamento, 13*(1), 342-353. <https://doi.org/10.18761/DH000166.set21>
- Arntzen, E., & Holth, P. (1997). Probability of stimulus equivalence as a function of training design. *The Psychological Record, 47*(2), 309–320.  
<https://doi.org/10.1007/BF03395227>
- Augoustinos, M., & Reynolds, K. J. (2001). Prejudice, racism, and Social Psychology (pp. 1-23). In: Augoustinos, M., & Reynolds, K. J. (Eds.). *Understanding prejudice, racism and social conflict*. Sage Publications.
- Auty, S., & Lewis, C. (2004). Exploring children's choice: The reminder effect of product placement. *Psychology & Marketing, 21*(9), 697-713.  
<https://doi.org/10.1002/mar.20025>
- Baer, D. M., Wolf, M. M., & Risley, T. R. (1968). Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis, 1*(1), 91–97.  
<https://doi.org/10.1901/jaba.1968.1-91>
- Baker, C. (2020, June). In a pandemic, all some people see is your color. *The Atlantic*.  
<https://www.theatlantic.com/magazine/archive/2020/06/coronavirus-racism/610609/>

Banaji, M. A., & Greenwald, A. G. (2013). *Blindspot: Hidden biases of good people*.

Delacorte Press.

Banks, K. H., Goswami, S., Goodwin, D., Petty, J., Bell, V., & Musa, I. (2021).

Interrupting internalized racial oppression: A community based ACT intervention. *Journal of Contextual Behavioral Science*, 20, 89-93.

<https://doi.org/10.1016/j.jcbs.2021.02.006>

Bareket, O., & Fiske, S. T. (2023). A systematic review of the ambivalent sexism literature:

Hostile sexism protects men's power; benevolent sexism guards traditional gender roles. *Psychological Bulletin*, 149(11-12), 637–698.

<https://doi.org/10.1037/bul0000400>

Barnes-Holmes, D., Barnes-Holmes, Y., & McEnteggart, C. (2020). Updating RFT (more

field than frame) and its implications for process-based therapy. *The Psychological Record*, 70(4), 605-624. <https://doi.org/10.1007/s40732-019-00372-3>

Barnes-Holmes, D., Barnes-Holmes, Y., Stewart, I., & Boles, S. (2010). A sketch of the

Implicit Relational Assessment Procedure (IRAP) and the Relational Elaboration and Coherence (REC) model. *The Psychological Record*, 60(4), 527–

542. <https://doi.org/10.1007/BF03395726>

Barnes-Holmes, D., & Harte, C. (2022). The IRAP as a measure of implicit cognition: A

case of Frankenstein's monster. *Perspectives on Behavior Science*, 45(3), 559-578.

<https://doi.org/10.1007/s40614-022-00352-z>

Barnes-Holmes, D., Murphy, C., Barnes-Holmes, Y., & Stewart, I. (2010). The Implicit

Relational Assessment Procedure: Exploring the impact of private versus public contexts and the response latency criterion on pro-White and anti-Black

- stereotyping among white Irish individuals. *The Psychological Record*, 60(1), 57–74. <https://doi.org/10.1007/BF03395695>
- Bar-Tal, D. (1996). Development of social categories and stereotypes in early childhood: The case of “the Arab” concept formation, stereotype and attitudes by Jewish children in Israel. *International journal of intercultural relations*, 20(3-4), 341-370. [https://doi.org/10.1016/0147-1767\(96\)00023-5](https://doi.org/10.1016/0147-1767(96)00023-5)
- Baron, A. S., & Banaji, M. R. (2006). The development of implicit attitudes: Evidence of race evaluations from ages 6 and 10 and adulthood. *Psychological Science*, 17(1), 53-58. <https://doi.org/10.1111/j.1467-9280.2005.01664.x>
- Baum, W. M. (2017). *Understanding Behaviorism: Behavior, Culture, and Evolution*. (3ed.). Wiley.
- Belisle, J., Dixon, M. R., & Stanley, C. R. (2023). Mapping relational networks of racial bias using Relational Density Theory. *The Psychological Record*, 73(4), 493–511. <https://doi.org/10.1007/s40732-023-00528-3>
- Belisle, J., Payne, A., & Paliliunas, D. (2022). A sociobehavioral model of racism against the Black community and avenues for anti-racism research. *Behavior Analysis in Practice*, 15, 1134-1150. <https://doi.org/10.1007/s40617-022-00702-y>
- Belisle, J., & Dixon, M. R. (2020). An exploratory analysis of relational density theory: Relational resistance and gravity. *Journal of Contextual Behavioral Science*, 16, 80-95. <https://doi.org/10.1007/s40614-020-00248-w>
- Bierley, C., McSweeney, F. K., & Vannieuwkerk, R. (1985). Classical conditioning of preferences for stimuli. *Journal of Consumer Research*, 12(3), 316–323. <https://doi.org/10.1086/208518>
- Bigler, R. S., & Liben, L. S. (2007). Developmental intergroup theory: Explaining and

- reducing children's social stereotyping and prejudice. *Current directions in Psychological Science*, 16(3), 162-166. <https://doi.org/10.1111/j.1467-8721.2007.00496.x>
- Blackson, E. A., Gerdes, M., Segan, E., Anokam, C., & Johnson, T. J. (2022). Racial bias toward children in the early childhood education setting. *Journal of Early Childhood Research*, 20(3), 277-292. <https://doi.org/10.1177/1476718X221087051>
- Bohner, G., & Dickel, N. (2011). Attitudes and attitude change. *Annual Review of Psychology*, 62, 391–417. <https://doi.org/10.1146/annurev.psych.121208.131609>
- Bortoloti, R., Pimentel, N., & Rose, J. C. D. (2014). Electrophysiological investigation of the functional overlap between semantic and equivalence relations. *Psychology & Neuroscience*, 7, 183-191. <https://doi.org/10.3922/j.psns.2014.015>
- Bortoloti, R., & Rose, J. C. D. (2007). Medida do grau de relacionamento entre estímulos equivalentes. *Psicologia: Reflexão e Crítica*, 20, 252-258. <https://doi.org/10.1590/S0102-79722007000200011>
- Bortoloti, R., & de Rose, J. C. (2009). Assessment of the relatedness of equivalent stimuli through a semantic differential. *The Psychological Record*, 59(4), 563–590. <https://doi.org/10.1007/BF03395682>
- Branton, R. P., & Jones, B. S. (2005). Reexamining racial attitudes: The conditional relationship between diversity and socioeconomic environment. *American Journal of Political Science*, 49(2), 359–372. <https://doi.org/10.1111/j.0092-5853.2005.00128.x>
- Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry*, 25(1), 49-59. [https://doi.org/10.1016/0005-7916\(94\)90063-9](https://doi.org/10.1016/0005-7916(94)90063-9)

- Callander, D., Newman, C. E., & Holt, M. (2015). Is sexual racism really racism? Distinguishing attitudes toward sexual racism and generic racism among gay and bisexual men. *Archives of Sexual Behavior, 44*(7), 1991–2000.  
<https://doi.org/10.1007/s10508-015-0487-3>
- Camilo, D. D. N., De Moura, G. B., Pimentel, C. E., Vera Noriega, J. A., & Gomes Cavalcanti, J. (2020). Preconceito racial entre crianças da educação infantil: revisitando Clark & Clark (1947). *CES Psicología, 13*(2), 32-45.  
<https://doi.org/10.21615/cesp.13.2.3>
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*(2), 81–105. <https://doi.org/10.1037/h0046016>
- Campón, R. R., & Carter, R. T. (2015). The relationship between racial discrimination and psychological distress: The role of racial identity and social support. *Journal of Counseling Psychology, 62*(2), 226–238. <https://doi.org/10.1037/cou0000065>
- Carifio, J. (1994). Sensitive data and students' tendencies to give socially desirable responses. *Journal of Alcohol & Drug Education, 39*(3), 74–84.
- Cartwright, A. R., Roche, B., O'Reilly, A., & Murphy, E. (2016). Using the Function Acquisition Speed Test (FAST) for assessing implicit gender bias. *The Psychological Record, 66*(2), 275-283. <https://doi.org/10.1007/s40732-016-0169-0>
- Carbado, D. W., Crenshaw, K. W., Mays, V. M., & Tomlinson, B. (2013). INTERSECTIONALITY: Mapping the Movements of a Theory1. *Du Bois review: social science research on race, 10*(2), 303-312.  
<https://doi.org/10.1017/S1742058X13000349>

- Carvalho, A. A. S., Mizael, T. M., & Sampaio, A. A. (2022). Racial prejudice and police stops: A systematic review of empirical literature. *Behavior analysis in practice, 15*(4), 1213–1220. <https://doi.org/10.1007/s40617-021-00578-4>
- Castelli, L., Zogmaister, C., & Tomelleri, S. (2009). The transmission of racial attitudes within the family. *Developmental Psychology, 45*(2), 586-591. <https://doi.org/10.1037/a0014619>
- Catagnus, R. M., Belisle, J., & Hantula, D. A. (2022). Behavior analysts' experiences discussing race and racism: Emotional responses and implications for training. *Behavior and Social Issues, 31*, e1051. <https://doi.org/10.1007/s42822-022-00054-7>
- Catania, C.A. (2017). Prejudice as verbally governed discrimination (pp. 254-262). In: *The ABCs of Behavior Analysis: Introduction to behavior and learning*. Sloan Publishing.
- Catania, C. (2018). *A behavior analytic perspective on discrimination and prejudice*. Chapter contributed to the 28th Congress of the Mexican Conference on the Analysis of Behavior. Veracruz, México. [https://www.researchgate.net/publication/334126626\\_A\\_behavior\\_analytic\\_perspective\\_on\\_discrimination\\_and\\_prejudice](https://www.researchgate.net/publication/334126626_A_behavior_analytic_perspective_on_discrimination_and_prejudice)
- Centers for Disease Control and Prevention. (2023). *Minority health: Racism and health*. <https://www.cdc.gov/minorityhealth/racism-disparities/index.html>
- Cesario, J. (2022). What can experimental studies of bias tell us about real-world group disparities? *Behavioral and Brain Sciences, 45*, e66, 1-71. <https://doi.org/10.1017/S0140525X21000017>

- Clark, K. B., & Clark, M. P. (1947). Racial identification and preference in negro children. In H. Proshansky & B. Seidenberg (Eds.), *Basic Studies in Social Psychology* (pp. 308-317). Holt Rinehart and Winston.
- Clark, R. (2004). Interethnic group and intraethnic group racism: Perceptions and coping in Black university students. *Journal of Black Psychology, 30*(4), 506-526.  
<https://doi.org/10.1177/0095798404268286>
- Cook, T. D., Campbell, D. T., & Shadish, W. (2002). *Experimental and quasi-experimental designs for generalized causal inference* (Vol. 1195). Boston, MA: Houghton Mifflin.
- Cooper, J. O., Heron, T. E., & Heward, W. L. (2020). *Applied behavior analysis* (3rd ed.). Pearson.
- Conrado, M., & Ribeiro, A. A. M. (2017). Homem negro, negro homem: Masculinidades e feminismo negro em debate. *Revista Estudos Feministas, 25*(1), 73–97.  
<https://doi.org/10.1590/1806-9584.2017v25n1p73>
- Corrington, A., Fa-Kaji, N. M., Hebl, M., Salgado, A., Brown, N. D., & Ng, L. (2023). The influence of social norms on the expression of anti-Black bias. *Journal of Business and Psychology, 38*(1), 89-108. <https://doi.org/10.1007/s10869-022-09822-2>
- Corneille, O., & Gawronski, B. (2024). Self-reports are better measurement instruments than implicit measures. *Nature Reviews Psychology, 1*-12.  
<https://doi.org/10.1038/s44159-024-00376-z>
- Correll, J., Park, B., Judd, C. M., & Wittenbrink, B. (2002). The police officer's dilemma: Using ethnicity to disambiguate potentially threatening individuals. *Journal of Personality and Social Psychology, 83*(6), 1314–1329. <https://doi.org/10.1037/0022-3514.83.6.1314>

- Correll, J., Park, B., Judd, C. M., & Wittenbrink, B. (2007). The influence of stereotypes on decisions to shoot. *European Journal of Social Psychology, 37*(6), 1102–1117. <https://doi.org/10.1002/ejsp.450>
- Correll, J., Urland, G. R., & Ito, T. A. (2006). Event-related potentials and the decision to shoot: The role of threat perception and cognitive control. *Journal of Experimental Social Psychology, 42*(1), 120–128. <https://doi.org/10.1016/j.jesp.2005.02.006>
- Correll, J., Wittenbrink, B., Crawford, M. T., & Sadler, M. S. (2014). Stereotypic vision: How stereotypes disambiguate visual stimuli. *Journal of Personality and Social Psychology, 108*(2), 219–233. <https://doi.org/10.1037/pspa0000015>
- Crano, W. D., & Prislin, R. (Eds.). (2011). *Attitudes and attitude change*. Psychology Press. <https://doi.org/10.4324/9780203838068>
- Crenshaw, K. (2013). Demarginalizing the intersection of race and sex: A black feminist critique of antidiscrimination doctrine, feminist theory and antiracist politics. In *Feminist legal theories* (pp. 23-51). Routledge.
- Critchfield, T. S. (2002). Evaluating the function of applied behavior analysis: A bibliometric analysis. *Journal of Applied Behavior Analysis, 35*(4), 423–426. <https://doi.org/10.1901/jaba.2002.35-423>
- Critchfield, T. S. (2011). Translational contributions of the experimental analysis of behavior. *The Behavior Analyst, 34*(1), 3-17. <https://doi.org/10.1007/BF03392227>
- Critchfield, T. S. (2011). To a young basic scientist, about to embark on a program of translational research. *The Behavior Analyst, 34*(2), 137-148. <https://doi.org/10.1007/BF03392245>
- Critchfield, T. S., & Perone, M. (1993). Verbal self-reports about matching to sample: Effects of the number of elements in a compound sample stimulus. *Journal of the*

*Experimental Analysis of Behavior*, 59(1), 193–214.

<https://doi.org/10.1901/jeab.1993.59-193>

Cummins, J., & Roche, B. (2020). Measuring differential nodal distance using the function acquisition speed test. *Behavioural Processes*, 178, 104179, 1-11.

<https://doi.org/10.1016/j.beproc.2020.104179>

Cummins, J., Tyndall, I., Curtis, A., & Roche, B. (2019). The Function Acquisition Speed Test (FAST) as a measure of verbal stimulus relations in the context of condom use. *The Psychological Record*, 69, 107–115. <https://doi.org/10.1007/s40732-016-0212-1>

Cummins, J., Roche, B., Tyndall, I., & Cartwright, A. (2018). The relationship between differential stimulus relatedness and implicit measure effect sizes. *Journal of the Experimental Analysis of Behavior*, 110(1), 24–38. <https://doi.org/10.1002/jeab.437>

Cvencek, D., Greenwald, A. G., & Meltzoff, A. N. (2016). Implicit measures for preschool children confirm self-esteem's role in maintaining a balanced identity. *Journal of Experimental Social Psychology*, 62, 50-57.

<https://doi.org/10.1016/j.jesp.2015.09.015>

Danis, A., Pêcheux, M. G., Lefèvre, C., Bourdais, C., & Serres-Ruel, J. (2008). A continuous performance task in preschool children: Relations between attention and performance. *European Journal of Developmental Psychology*, 5(4), 401-418.

<https://doi.org/10.1080/17405620600866800>

Danso, H. A., & Esses, V. M. (2001). Black experimenters and the intellectual test performance of White participants: The tables are turned. *Journal of Experimental Social Psychology*, 37(2), 158-165. doi:10.1006/jesp.2000.1444

- Dasgupta, N., & Greenwald, A. G. (2001). On the malleability of automatic attitudes: Combating automatic prejudice with images of admired and disliked individuals. *Journal of Personality and Social Psychology*, *81*(5), 800–814. <https://doi.org/10.1037/0022-3514.81.5.800>
- Dasgupta, N., McGhee, D. E., Greenwald, A. G., & Banaji, M. R. (2000). Automatic preference for White Americans: Eliminating the familiarity explanation. *Journal of Experimental Social Psychology*, *36*(3), 316-328. <https://doi.org/10.1006/jesp.1999.1418>
- de Carvalho, M. P., & de Rose, J. C. (2014). Understanding racial attitudes through the stimulus equivalence paradigm. *The Psychological Record*, *64*(3), 527-536. <https://doi.org/10.1007/s40732-014-0049-4>
- Dunham, Y., Baron, A. S., & Banaji, M. R. (2007). Children and social groups: A developmental analysis of implicit consistency in Hispanic Americans. *Self and Identity*, *6*(2-3), 238-255. <https://doi.org/10.1080/15298860601115344>
- De Houwer, J., Heider, N., Spruyt, A., Roets, A., & Hughes, S. (2015). The relational responding task. *Frontiers in Psychology*, *6*, 132367. <https://doi.org/10.3389/fpsyg.2015.00319>
- De Houwer, J., & Moors, A. (2012). How to define and examine implicit processes. *Perspectives on Psychological Science*, *7*(2), 231–237. <https://doi.org/10.1177/1745691611434367>
- De Houwer, J., Thomas, S., & Baeyens, F. (2001). Associative learning of likes and dislikes: A review of 25 years of research on human evaluative conditioning. *Psychological Bulletin*, *127*(6), 853–869. <https://doi.org/10.1037/0033-2909.127.6.853>

- de Melo, C. M., & Fernandes, D. M. (2025). Da " Questão do controle" à Questão Palestina: dando concretude à abstração skinneriana sobre controle social. *Perspectivas em Análise do Comportamento*, 16(2). <https://doi.org/10.18761/AB50CMMDM01>
- de Rose, J. C., & Bortoloti, R. (2007). A equivalência de estímulos como modelo do significado. *Acta Comportamental: Revista Latina de Análisis de Comportamiento*, 15(3), 83-102.  
<http://www.redalyc.org/articulo.oa?id=274520165006>
- Rose, J. C. D., McIlvane, W. J., Dube, W. V., Galpin, V. C., & Stoddard, L. T. (1988). Emergent simple discrimination established by indirect relation to differential consequences. *Journal of the Experimental Analysis of Behavior*, 50(1), 1-20.  
<https://doi.org/10.1901/jeab.1988.50-1>
- de Rose, J. C. (2022). Derived relations and meaning in responding to art. *Perspectives on Behavior Science*, 45(2), 445-455. <https://doi.org/10.1007/s40614-022-00334-1>
- de Rose, J. C., Hidalgo, M., & Vasconcellos, M. (2013). Controlling relations in baseline conditional discriminations as determinants of stimulus equivalence. *The Psychological Record*, 63(1), 85-98. <https://doi.org/10.11133/j.tpr.2013.63.1.007>
- de Souza Silva, T., Flandoli, B. R. G. X., & Mizael, T. M. (2022). Questões raciais na análise do comportamento: Uma análise preliminar sobre a baixa produção da área. *Perspectivas em análise do comportamento*, 13(1), 371-385.  
<https://doi.org/10.18761/VEEM.0077.out21>
- Dehon, E., Weiss, N., Jones, J., Faulconer, W., Hinton, E., & Sterling, S. (2017). A systematic review of the impact of physician implicit racial bias on clinical decision making. *Academic Emergency Medicine*, 24(8), 895–904.  
<https://doi.org/10.1111/acem.13214>

- Devine, P. G., Forscher, P. S., Austin, A. J., & Cox, W. T. (2012). Long-term reduction in implicit race bias: A prejudice habit-breaking intervention. *Journal of Experimental Social Psychology, 48*(6), 1267-1278. <https://doi.org/10.1016/j.jesp.2012.06.003>
- Dixon, J., Durrheim, K., & Tredoux, C. (2005). Beyond the Optimal Contact Strategy: A Reality Check for the Contact Hypothesis. *American Psychologist, 60*(7), 697–711. <https://doi.org/10.1037/0003-066X.60.7.697>
- Does, S., Ellemers, N., Dovidio, J. F., Norman, J. B., Mentovich, A., van der Lee, R., & Goff, P. A. (2018). Implications of research staff demographics for psychological science. *American Psychologist, 73*(5), 639–650. <https://doi.org/10.1037/amp0000199>
- Dougher, M. J., Augustson, E., Markham, M. R., Greenway, D. E., & Wulfert, E. (1994). The transfer of respondent eliciting and extinction functions through stimulus equivalence classes. *Journal of the Experimental Analysis of Behavior, 62*(3), 331–351. <https://doi.org/10.1901/jeab.1994.62-331>
- Dougher, M. J., Hamilton, D. A., Fink, B. C., & Harrington, J. (2007). Transformation of the discriminative and eliciting functions of generalized relational stimuli. *Journal of the experimental analysis of behavior, 88*(2), 179-197. <https://doi.org/10.1901/jeab.2007.45-05>
- Doorley, J. D., Goodman, F. R., Kelso, K. C., & Kashdan, T. B. (2020). Psychological flexibility: What we know, what we do not know, and what we think we know. *Social and Personality Psychology Compass, 14*(12), 1-11. <https://doi.org/10.1111/spc3.12566>

- dos Santos, S. L. D., & Rose, J. C. C. D. (2018). Investigating the impact of stimulus equivalence on children's food choice and preference. *Trends in Psychology, 26*(1), 1-14. <https://doi.org/10.9788/TP2018.1-01>
- dos Santos, S. L., & de Rose, J. C. (2019). Influence of cartoon characters on children's food preference via transfer of functions. *The Psychological Record, 69*(2), 153-163. <https://doi.org/10.1007/s40732-018-0327-7>
- Drake, C. E., O'Shea, D., & Barnes-Holmes, D. (2015). Measuring implicit racial evaluations using the IRAP: Reliability and convergent validity. *The Psychological Record, 65*(4), 681–698. \* <https://doi.org/10.1007/s40732-015-0131-1>
- Dunham, Y., Baron, A. S., & Banaji, M. R. (2007). Children and social groups: A developmental analysis of implicit consistency in Hispanic Americans. *Self and Identity, 6*(2-3), 238-255. <https://doi.org/10.1080/15298860601115344>
- Errasti, J., Martinez, H., Rodriguez, C., Marquez, J., Maldonado, A., & Mendez, A. (2019). Social context in a collective IRAP application about gender stereotypes: Mixed versus single gender groups. *The Psychological Record, 69*, 39–48. <https://doi.org/10.1007/s40732-018-0320-1>
- Fanon, F. (2020). *Pele negra, máscaras brancas* (1ª ed.). Ubu Editora.
- Farmer, N., Baginski, A., & Gordon, T. (2023, August 23). Behavioral Interventions to Reduce and Prevent Racial Bias. *Oxford Research Encyclopedia of Global Public Health*. <https://doi.org/10.1093/acrefore/9780190632366.013.475>
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G\*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191. <https://doi.org/10.3758/BF03193146>

- Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology*, *54*, 297–327. <https://doi.org/10.1146/annurev.psych.54.101601.145225>
- Fazio, R. H., Pietri, E. S., Rocklage, M. D., & Shook, N. J. (2015). Positive versus negative valence: Asymmetries in attitude formation and generalization as fundamental individual differences. In *Advances in experimental social psychology* (Vol. 51, pp. 97-146). Academic Press.
- Fernandes, S.C.S., Almeida, S.S.M., & Nascimento, C.B.S. (2008). Análise do preconceito racial em uma amostra de crianças brancas de 5 a 8 anos de idade. *PSICO*, Porto Alegre, *39*, (4), 441-447. Recuperado em: <https://dialnet.unirioja.es/servlet/articulo?codigo=5161537>
- Ferro, R., & Valero, L. (2008). Transfer of function through equivalence relations using pictures with strong emotional content. *European Journal of Behavior Analysis*, *9*(1), 13-27. <https://doi.org/10.1080/15021149.2008.11434292>
- Florez, J. A., Levin, M. E., & Hayes, S. C. (2019). Examining the relationship between psychological inflexibility and implicit bias. *Journal of Contextual Behavioral Science*, *12*, 1–5. <https://doi.org/10.1016/j.jcbs.2019.01.001>
- Flynn, M. K., Hernandez, J. O., Hebert, E. R., James, K. K., & Kusick, M. K. (2018). Cognitive fusion among Hispanic college students. *Journal of Contextual Behavioral Science*, *7*, 29–34. <https://doi.org/10.1016/j.jcbs.2017.11.003>
- Forscher, P. S., Lai, C. K., Axt, J. R., Ebersole, C. R., Herman, M., Devine, P. G., & Nosek, B. A. (2019). A meta-analysis of procedures to change implicit measures. *Journal of Personality and Social Psychology*, *117*(3), 522–559. <https://doi.org/10.1037/pspa0000160>

- Fox, D. J., & Jordan, V. B. (1973). Racial preference and identification of black, American Chinese, and white children. *Genetic Psychology Monographs*, 88(2), 229–286.
- França, D. X., & Monteiro, M. B. (2002). Identidade racial e preferência em crianças brasileiras de 5 a 10 anos. *Psicologia*, 16(2), 293-323. doi: <http://dx.doi.org/10.17575/rpsicol.v16i2.482>
- Gandarela, L., Boldrin, L. S., & Debert, P. (2020). Transfer of the avoidance function in equivalence classes. *The Psychological Record*, 70, 1–9. <https://doi.org/10.1007/s40732-019-00365-2>
- Gautier, T., & Droit-Volet, S. (2002). Attention and time estimation in 5- and 8-year-old children: A dual-task procedure. *Behavioural Processes*, 58(1-2), 57-66. [https://doi.org/10.1016/S0376-6357\(02\)00002-5](https://doi.org/10.1016/S0376-6357(02)00002-5)
- Gawronski, B. (2007). Attitudes can be measured! But what is an attitude?. *Social Cognition*, 25(5), 573-581. <https://doi.org/10.1521/soco.2007.25.5.573>
- Gingles, D., Watson-Thompson, J., Anderson-Carpenter, K. D., Tarbox, J., & Peterson, S. M. (2022). Leading the charge: A look inside the behavior analysis in practice emergency series of publications on systemic racism and police brutality. *Behavior Analysis in Practice*, 15(4), 1015-1022. <https://doi.org/10.1007/s40617-022-00759-9>
- Geist, T., Plezia, S., Cepeda-Benito, A. H., López, M., & Rodríguez Valverde, M. (2023). Online administration of the implicit relational assessment procedure: The Web-IRAP. *Psychological Record*, 73, 67–74. <https://doi.org/10.1007/s40732-022-00533-x>

- Glenn, S. S. (1988). Contingencies and metacontingencies: Toward a synthesis of behavior analysis and cultural materialism. *The Behavior Analyst, 11*(2), 161-179.  
<https://doi.org/10.1007/BF03392470>
- Glenn, S. S. (1991). Contingencies and metacontingencies: Relations among behavioral, cultural, and biological evolution. *Behavioral analysis of societies and cultural practices, 39-73*.
- Glick, P., & Fiske, S. T. (2001). An ambivalent alliance: Hostile and benevolent sexism as complementary justifications for gender inequality. *American Psychologist, 56*(2), 109–118. <https://doi.org/10.1037/0003-066X.56.2.109>
- Gonzalez, A. M., Steele, J. R., & Baron, A. S. (2017). Reducing children's implicit racial bias through exposure to positive out-group exemplars. *Child Development, 88*(1), 123-130. <https://doi.org/10.1111/cdev.12582>
- Gopal, D. P., Chetty, U., O'Donnell, P., Gajria, C., & Blackadder-Weinstein, J. (2021). Implicit bias in healthcare: Clinical practice, research and decision making. *Future healthcare journal, 8*(1), 40-48. <https://doi.org/10.7861/fhj.2020-0233>
- Graham, J. R., West, L. M., & Roemer, L. (2015). Valued living as a buffer against the impact of race-based stressors on mental health. *Journal of Contextual Behavioral Science, 4*(4), 268–275. <https://doi.org/10.1016/j.jcbs.2015.09.002>
- Greeley, A. M., & Sheatsley, P. B. (2017). Attitudes toward racial integration. In J. E. Parsons (Ed.), *Social Policy, and Public Policy* (pp. 241–250). Routledge.
- Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology, 74*(6), 1464–1480. <https://doi.org/10.1037/0022-3514.74.6.1464>

- Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology*, 97(1), 17–41. <https://doi.org/10.1037/a0015575>
- Greer, R. D. & Ross, D. E. (2008). *Verbal behavior analysis: inducing and expanding new verbal capabilities in children with language delays*. Pearson Education.
- Grey, I. M., & Barnes, D. (1996). Stimulus equivalence and attitudes. *The Psychological Record*, 46(2), 243.
- Guerin, B. (1994). *Analyzing social behavior: Behavior analysis and the social sciences*. Context Press.
- Guerin, B. (1994). Attitudes and beliefs as verbal behavior. *The Behavior Analyst*, 17(1), 155-163.  
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2733686/pdf/behavan00023-0157.pdf>
- Guerin, B. (2005). Combating everyday racial discrimination without assuming racists or racism: New intervention ideas from a contextual analysis. *Behavior and Social Issues*, 14, 46-70. <https://dx.doi.org/10.5210/bsi.v14i1.120>
- Guerin, B. (2003). Combating everyday racial discrimination without assuming racists or racism: New intervention ideas from a contextual behavioral analysis. *Behavior and Social Issues*, 12(1), 46–65. <https://doi.org/10.5210/bsi.v12i1.33>
- Harris, J. R. (2006). Parental selection: a third selection process in the evolution of human hairlessness and skin color. *Medical Hypotheses*, 66(6), 1053-1059.  
<https://doi.org/10.1016/j.mehy.2006.01.027>

- Harte, C., Barnes-Holmes, D., Moreira, M., de Almeida, J. H., Passarelli, D., & de Rose, J. C. (2021). Exploring a Training IRAP as a single participant context for analyzing reversed derived relations and persistent rule-following. *Journal of the Experimental Analysis of Behavior*, *115*(2), 460-480.
- Haydu, V. B., Gaça, L. B., Cognetti, N. P., Costa, C. E., & Tomanari, G. Y. (2015). Equivalência de estímulos e ciúme: Efeito de história pré-experimental. *Psicologia: Reflexão e Crítica*, *28*, 490-499. <https://doi.org/10.1590/1678-7153.201528308>
- Haydu, V. B., Camargo, J., & Bayer, H. (2015). Effects of preexperimental history on the formation of stimulus equivalence classes: A study with supporters of Brazilian soccer clubs. *Psychology & Neuroscience*, *8*(3), 385–396. <https://doi.org/10.1037/h0101276>
- Hayes, S. C., Barnes-Holmes, D., & Roche, B. (Eds.). (2001). *Relational Frame Theory: A post-Skinnerian account of human language and cognition*. Kluwer Academic/Plenum Publishers.
- Hayes, S. C. (2004). Acceptance and commitment therapy, relational frame theory, and the third wave of behavioral and cognitive therapies. *Behavior Therapy*, *35*(4), 639-665. [https://doi.org/10.1016/S0005-7894\(04\)80013-3](https://doi.org/10.1016/S0005-7894(04)80013-3)
- Hayes, S. C., Levin, M. E., Plumb-Villardaga, J., Villatte, J. L., & Pistorello, J. (2013). Acceptance and commitment therapy and contextual behavioral science. *Behavior Therapy*, *44*(2), 180–198. <https://doi.org/10.1016/j.beth.2009.08.002>
- Hayes, S. C., Luoma, J. B., Bond, F. W., Masuda, A., & Lillis, J. (2006). Acceptance and Commitment Therapy: Model, processes, and outcomes. *Behaviour Research and Therapy*, *44*(1), 1-25. <https://doi.org/10.1016/j.brat.2005.06.006>

- Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (2012). *Acceptance and Commitment Therapy: The Process and Practice of Mindful Change* (2nd ed.). Guilford Press.
- Hayes, S. C., Strosahl, K., Wilson, K. G., Bissett, R. T., Pistorello, J., Toarmino, D., & McCurry, S. M. (2004). Measuring experiential avoidance: A preliminary test of a working model. *The Psychological Record*, *54*(4), 553–578. <https://doi.org/10.1007/BF03395492>
- Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005). A meta-analysis on the correlation between the Implicit Association Test and explicit self-report measures. *Personality and Social Psychology Bulletin*, *31*(10), 1369–1385. <https://doi.org/10.1177/0146167205275613>
- Hofmann, W., De Houwer, J., Perugini, M., Baeyens, F., & Crombez, G. (2010). Evaluative conditioning in humans: A meta-analysis. *Psychological Bulletin*, *136*(3), 390–421. <https://doi.org/10.1037/a0018916>
- Hauserman, N., Walen, S. R., & Behling, M. (1973). Reinforced racial integration in the first grade: A study in generalization. *Journal of Applied Behavior Analysis*, *6*, 193–200. <https://dx.doi.org/10.1901/jaba.1973.6-193>
- Holmes, S. C., Zare, M., Haeny, A. M., & Williams, M. T. (2024). Racial stress and racial trauma. *Annual Review of Clinical Psychology*, *20*, 1–27. <https://doi.org/10.1146/annurev-clinpsy-081122-020235>
- Hugenberg, K., & Bodenhausen, G. V. (2004). Ambiguity in social categorization: The role of prejudice and facial affect in race categorization. *Psychological Science*, *15*(5), 342–345. <https://doi.org/10.1111/j.0956-7976.2004.00680.x>
- Hughes, S., & Barnes-Holmes, D. (2017). Relational frame theory: The basic account. In R. D. Zettle, S. C. Hayes, D. Barnes-Holmes, & A. Biglan (Eds.), *The Wiley Handbook*

*of Contextual Behavioral Science* (pp. 129–178). Wiley Blackwell.

<https://doi.org/10.1002/9781118489857.ch9>

Hughes, S., Barnes-Holmes, D., Van Dessel, P., de Almeida, J. H., Stewart, I., & De Houwer, J. (2018). On the symbolic generalization of likes and dislikes. *Journal of Experimental Social Psychology, 79*, 365-377.

<https://doi.org/10.1016/j.jesp.2018.09.002>

Hunter, M. (2007). The persistent problem of colorism: Skin tone, status, and inequality. *Sociology compass, 1*(1), 237-254. <https://doi.org/10.1111/j.1751-9020.2007.00006.x>

Hussey, I., Barnes-Holmes, D., & Murphy, C. (2015). The IRAP and racial bias: An examination of implicit racial responding. *The Psychological Record, 65*(3), 467–478. <https://doi.org/10.1007/s40732-015-0120-4>

Jacobsen, H. B., Klungsoyr, O., Landrø, N. I., Stiles, T. C., & Roche, B. T. (2020). MINDflex training for cognitive flexibility in chronic pain: A randomized, controlled cross-over trial. *Frontiers in Psychology, 11*, 604832, 1-16.

<https://doi.org/10.3389/fpsyg.2020.604832>

Jones, R. N. (2015). Practice and retest effects in longitudinal studies of cognitive functioning. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring, 1*(1), 101-102.

Jorm, A. F., Duncan-Jones, P., & Scott, R. (1989). An analysis of the re-test artefact in longitudinal studies of psychiatric symptoms and personality. *Psychological medicine, 19*(2), 487-493.

Johnston, J. M., & Pennypacker, H. S. (2009). Observing and recording. *Strategies and Tactics of Behavioral Research* 3rd Edition (pp. 115–138). New York: Routledge.

- Kalra, P. B., Gabrieli, J. D., & Finn, A. S. (2019). Evidence of stable individual differences in implicit learning. *Cognition*, *190*, 199-211.  
<https://doi.org/10.1016/j.cognition.2019.05.007>
- Katz, P. A., Sohn, M., & Zalk, S. R. (1975). Perceptual concomitants of racial attitudes in urban grade-school children. *Developmental Psychology*, *11*(2), 135–144. <https://doi.org/10.1037/h0076446>
- Kawakami, K., Phills, C. E., Steele, J. R., & Dovidio, J. F. (2007). (Close) distance makes the heart grow fonder: Improving implicit racial attitudes and interracial interactions through approach behaviors. *Journal of Personality and Social Psychology*, *92*(6), 957–971. <https://doi.org/10.1037/0022-3514.92.6.957>
- Kazdin, A. E. (2021). Single-case experimental designs: Characteristics, changes, and challenges. *Journal of the Experimental Analysis of Behavior*, *115*(1), 56-85.  
<https://doi.org/10.1002/jeab.638>
- Kohlenberg, B. S., Hayes, S. C., & Hayes, L. J. (1991). The transfer of contextual control over equivalence classes through equivalence classes: A possible model of social stereotyping. *Journal of the Experimental Analysis of Behavior*, *56*(3), 505-518.  
<https://doi.org/10.1901/jeab.1991.56-505>
- Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: A literature review. *Quality & Quantity*, *47*(4), 2025–2047. <https://doi.org/10.1007/s11135-011-9640-9>
- Kurdi, B., & Banaji, M. R. (2017). Re-evaluating evaluative conditioning: A meta-analysis of the formation of implicit attitudes. *Psychological Bulletin*, *143*(6), 601–637.  
<https://doi.org/10.1037/bul0000096>

- Lai, C. K., Marini, M., Lehr, S. A., Cerruti, C., Shin, J. E. L., Joy-Gaba, J. A., Ho, A. K., Teachman, B. A., Wojcik, S. P., Koleva, S. P., Frazier, R. S., Heiphetz, L., Chen, E. E., Turner, R. N., Haidt, J., Kesebir, S., Hawkins, C. B., Schaefer, H. S., Rubichi, S., Nosek, B. A. (2014). Reducing implicit racial preferences: I. A comparative investigation of 17 interventions. *Journal of Experimental Psychology: General*, *143*(4), 1765-1785. <https://doi.org/10.1037/a0036769>
- Lai, C. K., Skinner, A. L., Cooley, E., Murrar, S., Brauer, M., Devos, T., Calanchini, J., Xiao, Y. J., Pedram, C., Marshburn, C. K., Simon, S., Blanchar, J. C., Joy-Gaba, J. A., Conway, J., Redford, L., Klein, R. A., Roussos, G., Schellhaas, F. M. H., Burns, M., Nosek, B. A. (2016). Reducing implicit racial preferences: II. Intervention effectiveness across time. *Journal of Experimental Psychology: General*, *145*(8), 1001–1016. <https://doi.org/10.1037/xge0000179>
- Lalor, I. J., & Roche, B. (2025). Assessing the Predictive Validity of the Function Acquisition Speed Test in the Context of Voting Behavior. *Behavior and Social Issues*, 1-20. <https://doi.org/10.1007/s42822-025-00232-6>
- Landor, A. M. (2012). Does skin tone matter?: Exploring the impact of skin tone on colorism within families, racism, and racial socialization among African American adolescents (Doctoral dissertation, University of Georgia). <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3970169/>
- Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2005). International Affective Picture System (IAPS) [Database record]. APA PsycTests. <https://doi.org/10.1037/t66667-000>

- Lebrecht, S., Pierce, L. J., Tarr, M. J., & Tanaka, J. W. (2009). Perceptual other-race training reduces implicit racial bias. *PloS one*, 4(1), e4215.  
<https://doi.org/10.1371/journal.pone.0004215>
- Lerman, D. C., Tetreault, A., Hovanetz, A., Bellaci, E., Miller, J., Karp, H., ... & Toupard, A. (2010). Applying signal-detection theory to the study of observer accuracy and bias in behavioral assessment. *Journal of applied behavior analysis*, 43(2), 195-213.  
<https://doi.org/10.1901/jaba.2010.43-195>
- Levin, M. E., Luoma, J. B., Vilardaga, R., Lillis, J., Nobles, R., & Hayes, S. C. (2016). Examining the role of psychological inflexibility, perspective taking, and empathic concern in generalized prejudice. *Journal of Applied Social Psychology*, 46(3), 180-191. <https://doi.org/10.1111/jasp.12355>
- Levy, S. R., & Hughes, J. M. (2009). Development of racial and ethnic prejudice among children. In T. D. Nelson (Ed.), *Handbook of prejudice, stereotyping, and discrimination* (pp. 23-42). Psychology Press.
- Lillis, J., & Hayes, S. C. (2007). Applying acceptance, mindfulness, and values to the reduction of prejudice: A pilot study. *Behavior Modification*, 31(4), 389-411. <https://doi.org/10.1177/0145445506298413>
- Lima, M.E.O., dos Santos Rodrigues, H. & Santos, E.V. (2022). Sexual racism in Brazil: Aesthetic preference, beauty models and stereotypes. *Trends in Psychology*, 30, 480-496. <https://doi.org/10.1007/s43076-021-00128-5>
- Lindström, B., Selbing, I., Molapour, T., & Olsson, A. (2014). Racial bias shapes social reinforcement learning. *Psychological Science*, 25(3), 711-719.  
<https://doi.org/10.1177/0956797613514093>

- LoPresti, M. A., Brown, J. L., & Hoffman, M. A. (2022). Cumulative psychological effects of racism: PTSD and health outcomes. *Journal of Contextual Behavioral Science*, 23, 34–45. <https://doi.org/10.1016/j.jcbs.2022.02.004>
- Mace, F. C., & Critchfield, T. S. (2010). Translational research in behavior analysis: Historical traditions and imperative for the future. *Journal of the experimental analysis of behavior*, 93(3), 293-312. <https://doi.org/10.1901/jeab.2010.93-293>
- Macri, J. A., & Rogge, R. D. (2024). Examining domains of psychological flexibility and inflexibility as treatment mechanisms in acceptance and commitment therapy: A comprehensive systematic and meta-analytic review. *Clinical Psychology Review*, 110, 102432, 1-13. <https://doi.org/10.1016/j.cpr.2024.102432>
- Maddox, K. B. (2004). Perspectives on racial phenotypicality bias. *Personality and Social Psychology Review*, 8(4), 383-401. [https://doi.org/10.1207/s15327957pspr0804\\_4](https://doi.org/10.1207/s15327957pspr0804_4)
- Maina, I. W., Belton, T. D., Ginzberg, S., Singh, A., & Johnson, T. J. (2018). A decade of studying implicit racial/ethnic bias in healthcare providers using the implicit association test. *Social Science & Medicine*, 199, 219-229. <https://doi.org/10.1016/j.socscimed.2017.05.009>
- Martins, T. H. S., Rodrigues, R. M., Araújo, F. C. O., Cedro, A. M., Bortoloti, R., Varella, A. A. B., & Huziwara, E. M. (2023). Transfer of functions based on equivalence class formation using musical stimuli. *Journal of the Experimental Analysis of Behavior*, 120, 394-405. <https://doi.org/10.1002/jeab.881>
- Masuda, A., Hayes, S. C., Sackett, C. F., & Twohig, M. P. (2004). Cognitive defusion and self-relevant negative thoughts: Examining the impact of a ninety-year-old technique. *Behaviour Research and Therapy*, 42(4), 477-485. <https://doi.org/10.1016/j.brat.2003.10.008>

- Matsuda, K., Garcia, Y., Catagnus, R., & Brandt, J. A. (2020). Can behavior analysis help us understand and reduce racism? A review of the current literature. *Behavior Analysis in Practice, 13*(2), 336-347. <https://doi.org/10.1007/s40617-020-00411-4>
- McConahay, J. B. (1986). Modern racism, ambivalence, and the Modern Racism Scale. In J. F. Dovidio & S. L. Gaertner (Eds.), *Prejudice, discrimination, and racism* (pp. 91–125). Academic Press.
- Mekawi, Y., & Bresin, K. (2015). Is the evidence from racial bias shooting task studies a smoking gun? Results from a meta-analysis. *Journal of Experimental Social Psychology, 61*, 120–130. <https://doi.org/10.1016/j.jesp.2015.08.002>
- Miguel, C. F. (2016). Common and intraverbal bidirectional naming. *The Analysis of Verbal Behavior, 32*, 125-138. <https://doi.org/10.1007/s40616-016-0066-2>
- Mizael, T. M., Barrozo, S. C. V., & Hunziker, M. H. L. (2023). Uma interpretação analítico-comportamental da solidão da mulher negra [A behavior-analytic view of Black women's loneliness]. *Acta comportamentalia, 31*( 3), 403-418. <https://www.revistas.unam.mx/index.php/acom/article/view/86447>
- Mizael, T. M., de Almeida, J. H., Silveira, C. C., & de Rose, J. C. (2016). Changing racial bias by transfer of functions in equivalence classes. *The Psychological Record, 66*(3), 451-462. <https://doi.org/10.1007/s40732-016-0185-0>
- Mizael, T. M., de Almeida, J. H., Roche, B., & de Rose, J. C. (2021). Effectiveness of different training and testing parameters on the formation and maintenance of equivalence classes: Investigating prejudiced racial attitudes. *The Psychological Record, 71*(2), 265-277. <https://doi.org/10.1007/s40732-020-00435-w>
- Mizael, T. M., de Castro, M. S. L. B., & Dittrich, A. (2021). Uma interpretação analítico-comportamental do colorismo e de suas implicações clínicas. *Acta*

*Comportamentalia: Revista Latina de Análisis de Comportamiento*, 29(4), 65-81.

Retrieved from

<https://www.redalyc.org/journal/2745/274571372005/274571372005>

Mizael, T. M., & de Rose, J. C. (2017). *Behavior analysis and racial prejudice:*

*Interpretative possibilities and challenges* [Análise do comportamento e preconceito

racial: Possibilidades de interpretação e desafios]. *Acta Comportamentalia*, 25(3),

365-377. <https://doi.org/10.32870/ac.v25i3.61632>

Mizael, T. M., dos Santos, S. L., & de Rose, J. C. C. (2016). Contribuições do paradigma

de equivalência de estímulos para o estudo das atitudes. *Interação em Psicologia*,

20(2), 124-134. Retrieved from

<https://revistas.ufpr.br/psicologia/article/download/46278/29997>

Mizael, T. M., & Sampaio, A. A. (2019). Racismo institucional: Aspectos comportamentais e culturais da abordagem policial. *Acta comportamentalia*, 27(2).

<https://doi.org/10.32870/ac.v27i2.69861>

Moreira, A. (2019). O que é racismo recreativo? [What is recreational racism?]. *Pólen*.

Moxon, P. D., Keenan, M., & Hine, L. (1993). Gender-role stereotyping and stimulus equivalence. *The Psychological Record*, 43(3), 381.

Moya, M., Glick, P., Expósito, F., De Lemus, S., & Hart, J. (2007). It's for your own good:

Benevolent sexism and women's reactions to protectively justified

restrictions. *Personality and Social Psychology Bulletin*, 33(10), 1421-1434.

<https://doi.org/10.1177/0146167207304790>

Munanga, K. (2004). Uma abordagem conceitual das noções de raça, racismo, identidade e etnia. [A conceptual approach of the concepts of race, racism, identity, and

- ethnicity]. In. Brandão, A. A. P. (Org.). *Programa de Educação sobre o Negro na Sociedade Brasileira*. EdUFF.
- Nesdale, D. (1999). Social identity and ethnic prejudice in children. *Psychology and society*, 92-110.
- Nesdale, D., Maass, A., Durkin, K., & Griffiths, J. (2005). Group norms, threat, and children's racial prejudice. *Child Development*, 76(3), 652-663.  
<https://doi.org/10.1111/j.1467-8624.2005.00869.x>
- Nevin, J. A., Mandell, C., & Atak, J. R. (1983). The analysis of behavioral momentum. *Journal of the Experimental Analysis of Behavior*, 39(1), 49-59.  
<https://doi.org/10.1901/jeab.1983.39-49>
- Nevin, J. A. (1992). An integrative model for the study of behavioral momentum. *Journal of the experimental analysis of behavior*, 57(3), 301-316.  
<https://doi.org/10.1901/jeab.1992.57-301>
- Nevin, J. A., & Grace, R. C. (2000). Behavioral momentum and the law of effect. *Behavioral and Brain Sciences*, 23(1), 73–90. <https://doi.org/10.1017/S0140525X00002405>
- Nevin, J. A., & Shahan, T. A. (2011). Behavioral momentum theory: Equations and applications. *Journal of Applied Behavior Analysis*, 44(4), 877-895.  
<https://doi.org/10.1901/jaba.2011.44-877>
- Nogueira, O. (2007). Preconceito racial de marca e preconceito racial de origem: Sugestão de um quadro de referência para a interpretação do material sobre relações raciais no Brasil. *Tempo Social*, 19(1), 287–308. <https://doi.org/10.1590/s0103-20702007000100015>

- Nosek, B. A., Banaji, M. R., & Greenwald, A. G. (2002). Harvesting implicit group attitudes and beliefs from a demonstration web site. *Group Dynamics: Theory, Research, and Practice*, 6(1), 101–115. <https://doi.org/10.1037/1089-2699.6.1.101>
- Nosek, B. A., Smyth, F. L., Hansen, J. J., Devos, T., Lindner, N. M., Ratliff (Ranganath), K. A., Smith, C. T., Olson, K. R., Chugh, D., Greenwald, A. G., & Banaji, M. R. (2007). Pervasiveness and correlates of implicit attitudes and stereotypes. *European Review of Social Psychology*, 18, 36-88. <https://doi.org/10.1080/10463280701489053>
- Nosek, B. A. (2007). Implicit–explicit relations. *Current Directions in Psychological Science*, 16(2), 65–69. <https://doi.org/10.1111/j.1467-8721.2007.00477.x>
- Nosek, B. A., & Smyth, F. L. (2007). A multitrait-multimethod validation of the implicit association test. *Experimental psychology*, 54(1), 14-29. <https://doi.org/10.1027/1618-3169.54.1.14>
- Oliver, J. E., & Mendelberg, T. (2000). Reconsidering the environmental determinants of white racial attitudes. *American Journal of Political Science*, 574-589. <https://doi.org/10.2307/2669265>
- Olson, M. A., & Fazio, R. H. (2001). Implicit attitude formation through classical conditioning. *Psychological Science*, 12(5), 413–417. <https://doi.org/10.1111/1467-9280.00376>
- Olson, M. A., & Fazio, R. H. (2006). Reducing automatically activated racial prejudice through implicit evaluative conditioning. *Personality and Social Psychology Bulletin*, 32(4), 421–433. <https://doi.org/10.1177/0146167205282734>
- Osgood, C. E., Suci, G. J., & Tannenbaum, P. H. (1957). *The measurement of meaning* (No. 47). University of Illinois press.

- Oswald, F. L., Mitchell, G., Blanton, H., Jaccard, J., & Tetlock, P. E. (2013). Predicting ethnic and racial discrimination: A meta-analysis of IAT criterion studies. *Journal of Personality and Social Psychology, 105*(2), 171–192. <https://doi.org/10.1037/a0032734>
- O'Reilly, A., Roche, B., Ruiz, M., Tyndall, I., & Gannon, S. (2012). The Function Acquisition Speed Test (FAST): A behavior analytic implicit test for assessing stimulus relations. *The Psychological Record, 62*(3), 507–528. <https://doi.org/10.1007/BF03395815>
- Otta, E., Leme, M. A. V. S., Lima, M. P., & Sampaio, S. M. R. (1983). Profecias auto-realizadoras em sala de aula: expectativas de estudantes de psicologia como determinantes não-intencionais do desempenho [Self-fulfilling prophecies in the classroom: expectations of psychology students as unintentional determinants of behavior]. *Psicologia, 9*, 27-42.
- O'Reilly, A., Roche, B., Gavin, A., & Ruiz, M. R. (2013). A function acquisition speed test for equivalence relations (FASTER). *The Psychological Record, 63*, 707–724. <https://doi.org/10.11133/j.tpr.2013.63.4.001>
- O'Reilly, A., Roche, B., Ruiz, M., Tyndall, I., & Gavin, A. (2012). The function acquisition speed test (FAST): a behavior analytic implicit test for assessing stimulus relations. *The Psychological Record, 62*(3), 507–528. <https://doi.org/10.1007/BF03395817>
- O'Reilly, A. G., Roche, B., & Cartwright, A. (2015). Function over form: A behavioral approach to implicit attitudes. In *Exploring implicit cognition: learning, memory, and social cognitive processes* (pp. 162-182). IGI Global.

- Pacheco, A. C. L. (2008). Branca para casar, mulata para f..., negra para trabalhar: escolhas afetivas e significados de solidão entre mulheres negras em Salvador, Bahia. *Campinas: Universidade Estadual de Campinas*.
- Paige, B., Askew, L., Haakmat, N., Otañez, L., & Monk-Turner, E. (2024). Public perceptions of justification in the deadly police shootings of Black men: A content analysis of social media posts in 2016 and 2020. *Journal of Ethnicity in Criminal Justice*, 22(2), 140–162. <https://doi.org/10.1080/15377938.2024.2342796>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
- Parks, G. S., & Rachlinski, J. J. (2009). Implicit bias, election'08, and the myth of a post-racial America. *Fla. St. UL Rev.*, 37, 659.
- Passarelli, D. A., Amd, M., de Oliveira, M. A., & de Rose, J. C. (2022). Augmenting salivation, but not evaluations, through subliminal conditioning of eating-related words. *Behavioural Processes*, 194, 104541. <https://doi.org/10.1016/j.beproc.2021.104541>
- Passarelli, D. A., Rico, A. S., & Silvestre, M. H. (2023). *Contributions of behavior analysis to the understanding of racism* [Contribuições da análise do comportamento para a compreensão do racismo]. *Revista Brasileira de Análise do Comportamento*, 19(1), 33–44. <https://doi.org/10.18542/rebac.v19i1.14941>
- Passarelli, D. A., Roche, B., Mizael, T. M., Silvestre, M. H., & de Rose, J. C. (2024). Reducing racial bias using a conflicting relations paradigm: A systematic replication of Mizael et al. (2016). *Behavior and Social Issues*, 1-28. <https://doi.org/10.1007/s42822-024-00181-6>

- Passarelli, D. A., Mizael, T. M., Silvestre, M. H., & de Rose, J. C. (2024). Examining implicit and explicit anti-Black racial bias in children: A repeated measure design using the Function Acquisition Speed Test (FAST). *European Journal of Behavior Analysis*, 1-21. <https://doi.org/10.1080/15021149.2024.2388487>
- Payne, B. K. (2001). Prejudice and perception: The role of automatic and controlled processes in misperceiving a weapon. *Journal of Personality and Social Psychology*, 81(2), 181–192. <https://doi.org/10.1037/0022-3514.81.2.181>
- Payne, B. K., Burkley, M. A., & Stokes, M. B. (2008). Why do implicit and explicit attitude tests diverge? The role of structural fit. *Journal of Personality and Social Psychology*, 94(1), 16–31. <https://doi.org/10.1037/0022-3514.94.1.16>
- Payne, B. K., & Correll, J. (2020). Race, weapons, and the perception of threat. *Advances in Experimental Social Psychology*, 62, 1–40. <https://doi.org/10.1016/bs.aesp.2020.04.001>
- Peck, T. C., Seinfeld, S., Aglioti, S. M., & Slater, M. (2013). Putting yourself in the skin of a black avatar reduces implicit racial bias. *Consciousness and cognition*, 22(3), 779-787. <https://doi.org/10.1016/j.concog.2013.04.016>
- Phelan, J. E., Sanchez, D. T., & Broccoli, T. L. (2010). The danger in sexism: The links among fear of crime, benevolent sexism, and well-being. *Sex Roles*, 62(1), 35-47. <https://doi.org/10.1007/s11199-009-9711-8>
- Pierce, J. W., Gray, J. R., Simpson, S., MacAskill, M. R., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, 51(1), 195–203. <https://doi.org/10.3758/s13428-018-01193-y>.

- Pierre, T. S., White, K. S., & Johnson, E. K. (2023). Who is running our experiments? The influence of experimenter identity in the marshmallow task. *Cognitive Development, 65*, 101271. <https://doi.org/10.1016/j.cogdev.2022.101271>
- Power, P. M., Barnes-Holmes, D., Hussey, I., & Murphy, C. (2017a). An electrophysiological analysis of racial bias using the IRAP. *The Psychological Record, 67*(2), 185–200. <https://doi.org/10.1007/s40732-016-0210-6>
- Power, P. M., Hussey, I., & Barnes-Holmes, D. (2017b). Implicit racial attitudes among Black and White participants: An IRAP investigation. *The Psychological Record, 67*(3), 329–342. <https://doi.org/10.1007/s40732-017-0245-3>
- Ribeiro, K. L., Magalhães, T. S., de Souza Costa, M. D., Oliveira, Y. N., de Oliveira Henklain, M. H., & Garcia, R. V. B. (2020). Equivalência de estímulos elucidada como propagandas influenciam o comportamento alimentar infantil. *Revista Brasileira de Terapia Comportamental e Cognitiva, 22*. <https://doi.org/10.31505/rbtcc.v22i1.1422>
- Rice, G. E. (1964). Prejudice and competitive behavior: The role of education and race. *The Psychological Record, 14*(4), 489–498. \* <https://doi.org/10.1007/BF03393552>
- Ryle, G., & Tanney, J. (2009). *The concept of mind*. Routledge.
- Roche, B., & Barnes-Holmes, D. (2003). Behavior analysis and social constructionism: Some points of contact and departure. *The Behavior Analyst, 26*(2), 215-231. <https://doi.org/10.1007/BF03392077>
- Roche, B., Passarelli, D. A., & Wall, M. (2025). Assessing the viability of the Function Acquisition Speed Test (FAST) method as a social research tool for indexing socially sensitive verbal relations. *The Psychological Record, 1-14*. <https://doi.org/10.1007/s40732-025-00660-1>

- Rodrigues, W. C., Marcelino, M. R., Arantes, A., & Chamel, N. E. (2022). Preconceito racial entre universitários: implicações sobre o uso do FAST no estudo da aprendizagem verbal. *Perspectivas em Análise do Comportamento*, 13(2), 054-069. <https://doi.org/10.18761/PAC000774.fev22>
- Rogers, R. W., & Prentice-Dunn, S. (1981). Deindividuation and anger-mediated interracial aggression: Unmasking regressive racism. *Journal of Personality and Social Psychology*, 41(1), 63–73. <https://doi.org/10.1037/0022-3514.41.1.63>
- Romani, P. W., Hays, T. M., Glover, J., & Swain-Cockrell, A. (2024). Training healthcare professionals to discriminate between examples and non-examples of racial microaggressions. *Behavioral Interventions*, 39(3), e2013, 1-13. <https://doi.org/10.1002/bin.2013>
- Ross, H. J. (2015). *Everyday bias: Identifying and navigating unconscious judgments in our daily lives*. Rowman & Littlefield.
- Rosenberg, M. J., & Hovland, C. I. (1960). Cognitive, Affective, and Behavioral Components of Attitude. In M. J. Rosenberg, C. I. Hovland, W. J. McGuire, R. P. Abelson, & J. W. Brehm (Eds.), *Attitude Organization and Change: An Analysis of Consistency among Attitude Components* (pp. 1-14). New Haven, CT: Yale University Press.
- Saini, V., & Vance, H. (2020). Systemic racism and cultural selection: A preliminary analysis of metacontingencies. *Behavior and Social Issues*, 29(1), 52-63. <https://doi.org/10.1007/s42822-020-00040-0>
- Sagonas, C., Antonakos, E., Tzimiropoulos, G., Zafeiriou, S., & Pantic, M. (2016). 300 Faces in-the-wild challenge: Database and results. *Image and Vision Computing*, 47, 3–18. <https://doi.org/10.1016/j.imavis.2016.01.002>

- Santos, W. S. D., Gouveia, V. V., Navas, M. S., Pimentel, C. E., & Gusmão, E. É. D. S. (2006). Escala de racismo moderno: adaptação ao contexto brasileiro. *Psicologia em estudo, 11*, 637-645. <https://doi.org/10.1590/S1413-73722006000300020>
- Sawyer, J., & Gampa, A. (2018). Implicit and explicit racial attitudes changed during Black Lives Matter. *Personality and Social Psychology Bulletin, 44*(7), 1039-1059. <https://doi.org/10.1177/0146167218757454>
- Schneider, D. (2005). *The Psychology of stereotyping*. Guilford Press.
- Schnierle, J., Christian-Brathwaite, N., & Louisias, M. (2019). Implicit bias: what every pediatrician should know about the effect of bias on health and future directions. *Current problems in pediatric and adolescent health care, 49*(2), 34-44. <https://doi.org/10.1016/j.cppeds.2019.01.003>
- Schultz, P. W., Shriver, C., Tabanico, J. J., & Khazian, A. M. (2004). Implicit connections with nature. *Journal of environmental psychology, 24*(1), 31-42.
- Seaton, E. K., & Iida, M. (2019). Racial discrimination and racial identity: Daily moderation among Black youth. *American Psychologist, 74*(1), 117–127. <https://doi.org/10.1037/amp0000367>
- Setoh, P., Lee, K. J., Zhang, L., Qian, M. K., Quinn, P. C., Heyman, G. D., & Lee, K. (2019). Racial categorization predicts implicit racial bias in preschool children. *Child Development, 90*(1), 162-179. <https://doi.org/10.1111/cdev.12851>
- Sinclair, S., Dunn, E., & Lowery, B. (2005). The relationship between parental racial attitudes and children's implicit prejudice. *Journal of Experimental Social Psychology, 41*(3), 283-289. <https://doi.org/10.1016/j.jesp.2004.06.003>
- Sidman, M. (1994). *Equivalence relations and behavior: A research story*. Authors Cooperative.

- Sidman, M., & Tailby, W. (1982). Conditional discrimination vs. matching to sample: An expansion of the test paradigm. *Journal of the Experimental Analysis of Behavior*, 37(1), 5-22. <https://dx.doi.org/10.1901/jeab.1982.37-5>
- Sidman, M. (2000). Equivalence relations and the reinforcement contingency. *Journal of the Experimental Analysis of behavior*, 74(1), 127-146.  
<https://doi.org/10.1901/jeab.2000.74-127>
- Silva, G. M., & Leão, L. T. S. (2012). O paradoxo da mistura: Identidades, desigualdades e percepção de discriminação entre brasileiros pardos. *Revista Brasileira de Ciências Sociais*, 27(80), 117-133. <https://doi.org/10.1590/S010269092012000300007>
- Silva, T. de S., Flandoli, B. R. G. X., & Mizael, T. M. (2022). Questões raciais na análise do comportamento: Uma análise preliminar sobre a baixa produção da área [Racial issues in behavior analysis: A preliminary analysis of the few studies published on the topic]. *Perspectivas Em Análise Do Comportamento*, 13(1), 371–385.  
<https://doi.org/10.18761/VEEM.0077.out21>
- Silveira, M. V., Aggio, N. M., Cortez, M. D., Bortoloti, R., Rico, V. V., & de Rose, J. C. (2016). Maintenance of equivalence classes and transfer of functions: The role of the nature of stimuli. *The Psychological Record*, 66(1), 65-74.
- Silveira, M. V., Mackay, H. A., & de Rose, J. C. (2018). Measuring the “transfer of meaning” through members of equivalence classes merged via a class-specific reinforcement procedure. *Learning & Behavior*, 46(2), 157-170.  
<https://doi.org/10.3758/s13420-017-0298-6>
- Slobodin, O., Noufi, T., & Caspi, Y. (2023). Interventions to Reduce Racial Bias and Prejudice in Preschool Children: A Systematic Review of the Literature Since

2010. *School Mental Health*, 15(4), 1017-1032. <https://doi.org/10.1007/s12310-023-09608-8>
- Skinner, B. F. (1957). *Verbal behavior*. New York: Appleton-Century-Crofts.
- Skinner, B. F. (1953). *Science and human behavior*. Macmillan.
- Smith, A. D. (1995). *The ethnic origins of nations*. Blackwell.
- Smeets, P. M., & Barnes-Holmes, D. (2003). Children's emergent preferences for soft drinks: Stimulus-equivalence and transfer. *Journal of Economic Psychology*, 24(5), 603–618.
- Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization 1. *Journal of applied behavior analysis*, 10(2), 349-367. <https://doi.org/10.1901/jaba.1977.10-349>
- Suarez, V. D., Najdowski, A. C., Persicke, A., & Tarbox, J. (2024). A laboratory evaluation of the effects of empathy training on racial bias. *The Psychological Record*, 74(3), 271-289. <https://doi.org/10.1007/s40732-024-00606-z>
- Strand, R. C., & Arntzen, E. (2020). Social categorization and stimulus equivalence: A systematic replication. *The Psychological Record*, 70(1), 47-63. <https://doi.org/10.1007/s40732-019-00364-3>
- Tappan, M. B. (2006). Reframing internalized oppression and internalized domination: From the psychological to the sociocultural. *Teacher College Records*, 108(10), 2115-2144. <https://doi.org/10.1111/j.1467-9620.2006.00776.x>
- Tajeu, G. S., Halanych, J., Juarez, L., Stone, J., Stepanikova, I., Green, A., & Cherrington, A. L. (2018). Exploring the association of healthcare worker race and occupation with implicit and explicit racial bias. *Journal of the National Medical Association*, 110(5), 464-472. <https://doi.org/10.1016/j.jnma.2017.12.001>

- Telles, E. E. (2014). Race in another America: The significance of skin color in Brazil. In *Race in another America*. Princeton University Press.
- Torrez, B., Dupree, C. H., & Kraus, M. W. (2024). How race influences perceptions of objectivity and hiring preferences. *Journal of Experimental Social Psychology, 110*, 104524. <https://doi.org/10.1016/j.jesp.2023.104524>
- Van de Mortel, T. F. (2008). Faking it: social desirability response bias in self-report research. *Australian Journal of Advanced Nursing, The, 25*(4), 40-48
- Wallace, B.W. (2003). *Match to Sample Program III* (Software de Computador, Versão III). Umass / Eunice Kennedy Shriver Center's Behavioral Sciences Department. Recuperado de: <http://mts3software.com/>
- Watt, A., Keenan, M., Barnes, D., & Cairns, E. (1991). Social categorization and stimulus equivalence. *The Psychological Record, 41*, 33-50.
- Watters, A., Cummins, J., & Roche, B. (2023). How to build and how not to build an implicit measure in behavior analysis: A case study using the Function Acquisition Speed Test. *Perspectives on Behavior Science, 46*(3), 459-492. <https://doi.org/10.1007/s40614-023-00387-w>
- West, L. M., Graham, J. R., & Roemer, L. (2013). Functioning in the face of racism: Preliminary findings on the buffering role of values clarification in a Black American sample. *Journal of Contextual Behavioral Science, 2*(1-2), 1-8. <https://doi.org/10.1016/j.jcbs.2013.03.003>
- Williams, M. T., Kanter, J. W., Peña, A., Ching, T. H., & Oshin, L. (2020). Reducing microaggressions and promoting interracial connection: The racial harmony workshop. *Journal of Contextual Behavioral Science, 16*, 153-161. <https://doi.org/10.1016/j.jcbs.2020.04.008> .

- Wolf, M. M. (1978). Social validity: the case for subjective measurement or how applied behavior analysis is finding its heart 1. *Journal of applied behavior analysis*, *11*(2), 203-214. <https://doi.org/10.1901/jaba.1978.11-203>
- Zanon, R., De Houwer, J., Gast, A., & Smith, C. T. (2014). When does relational information influence evaluative conditioning? *Quarterly Journal of Experimental Psychology*, *67*(11), 2105-2122. <https://doi.org/10.1080/17470218.2014.907324>
- Žeželj, I., Jakšić, I., & Jošić, S. (2015). How contact shapes implicit and explicit preferences: attitudes toward Roma children in inclusive and non-inclusive environment. *Journal of Applied Social Psychology*, *45*(5), 263-273.

## APÊNDICES

### Apêndice I

#### *Estratégia de Busca Da Revisão Sistemática*

(Any Field: "Racial Bias" OR Any Field: Racism OR Any Field: Prejudice) AND Journal Title: {Journal of the Experimental Analysis of Behavior} OR {Corrective & Social Psychiatry & Journal of Applied Behavior Therapy} OR {Journal of Applied Behavior} OR {Journal of Applied Behavior Analysis} OR {Journal of Applied Behavioral Science} OR {Behavior and Social Issues} OR {Behavior Analysis in Practice} OR {European Journal of Behavior Analysis} OR {Behavior Analysis & Social Action} OR {Behavior Analysis: Research and Practice} OR {Perspectives on Behavior Science} OR {Behavior Modification} OR {Behavioral Interventions} OR {Learning & Behavior} OR {Behavioral and Brain Sciences} OR {Journal of Behavioral Education} OR {Journal of Organizational Behavior Management} OR {Analysis of Verbal Behavior} OR {Journal of Contextual Behavioral Science} OR {International Journal of Behavioral Consultation and Therapy} OR {International Journal of Behavioral Development} OR {International Journal of Behavioral Medicine} OR {International Journal of Behavioral Sciences} OR {Behavioural Processes} OR {Behavioural Analysis & Modification} OR {European Journal of Behavioural Analysis & Modification} OR {International Behavioural Scientist} on Databases APA PsycINFO AND APA PsycArticles

## Apêndice II

### Lista de Siglas e Abreviações

DDS – Discrimination and Diversity Scale – Escala de Discriminação e Diversidade; mede atitudes e percepções sobre diversidade e discriminação.

MMRS – Modified Modern Racism Scale – Escala de Racismo Moderno Modificada; avalia formas sutis e contemporâneas de racismo.

IRAP – Implicit Relational Assessment Procedure – Procedimento de Avaliação Relacional Implícita; mede padrões de respostas relacionais sob pressão temporal.

SRE / SRE-R – Schedule of Racist Events – Inventário de Eventos Racistas; avalia a frequência e o impacto de experiências de racismo (ano atual ou passado).

VVIQ – Vividness of Visual Imagery Questionnaire – Questionário de Vividez da Imaginação Visual; mede a clareza de imagens mentais.

PANAS – Positive and Negative Affect Schedule – Escala de Afeto Positivo e Negativo; avalia estados afetivos positivos e negativos.

SUDS – Subjective Units of Distress Scale – Unidades Subjetivas de Desconforto; mede a intensidade subjetiva de sofrimento emocional.

SDS – Semantic Differential Scale – Escala de Diferencial Semântico; avalia significados atribuídos a estímulos em dimensões avaliativas.

SDO – Social Dominance Orientation – Orientação à Dominância Social; mede a preferência por hierarquias entre grupos sociais.

VLQ – Valued Living Questionnaire – Questionário de Vida com Valores; avalia a consistência entre ações e valores pessoais.

DASS-21 – Depression Anxiety Stress Scales – Escalas de Depressão, Ansiedade e Estresse; avalia sintomas emocionais.

SAM – Self-Assessment Manikin – Autoavaliação por Manikin; mede valência emocional, ativação e dominância.

EEG – Electroencephalography – Eletroencefalografia; registra a atividade elétrica cerebral.

MLQ-P – Meaning in Life Questionnaire – Presence – Questionário de Sentido da Vida (Presença); avalia a percepção de significado na vida.

PAB – Prejudice Against Blacks Scale – Escala de Preconceito contra Pessoas Negras; avalia atitudes preconceituosas explícitas.

PVQ-RR-ST – Portrait Values Questionnaire – Revised – Short Term – Questionário de Valores Pessoais (versão breve); avalia valores humanos básicos.

VQ-O – Valuing Questionnaire – Obstruction – Questionário de Valorização (Obstrução); mede barreiras à ação orientada por valores.

AS – Allophilia Scale – Escala de Allophilia; avalia atitudes positivas em relação a grupos externos.

IAA-6 – Intergroup Anxiety and Avoidance – Ansiedade e Evitação Intergrupais; mede desconforto em interações intergrupais.

MEIM-12 – Multigroup Ethnic Identity Measure – Medida de Identidade Étnica; avalia identificação e pertencimento étnico.

CoBRAS – Color-Blind Racial Attitudes Scale – Escala de Atitudes Raciais Daltônicas; mede a negação das desigualdades raciais.

SR2K – Symbolic Racism Scale – Escala de Racismo Simbólico; avalia formas indiretas de racismo.

CCAS – Color-Conscious Awareness Scale – Escala de Consciência Racial; avalia o reconhecimento da raça e de seus efeitos sociais.

AAQ-II – Acceptance and Action Questionnaire – Questionário de Aceitação e Ação; mede flexibilidade psicológica.

FAST – Function Acquisition Speed Test – Teste de Velocidade de Aquisição Funcional; avalia a rapidez na aquisição de classes funcionais de resposta.

DT – Doll Test – Teste da Boneca; avalia preferências e atribuições raciais.

RT – Relational Test – Teste Relacional; tarefas relacionais que avaliam a frequência com que faces negras e brancas são associadas a símbolos positivos ou negativos.

FT – Feeling Thermometer – Termômetro de Sentimentos; mede avaliação afetiva contínua em relação a grupos.

ABA – Applied Behavior Analysis (Análise do Comportamento Aplicada)

ABAI – Association for Behavior Analysis International

ACT – Acceptance and Commitment Therapy (Terapia de Aceitação e Compromisso)

CAPES – Coordenação de Aperfeiçoamento de Pessoal de Nível Superior

CFP – Conselho Federal de Psicologia

CNPq – Conselho Nacional de Desenvolvimento Científico e Tecnológico

CRP – Conselho Regional de Psicologia

CS-US – Conditioned Stimulus–Unconditioned Stimulus (Estímulo Condicionado–Estímulo Incondicionado)

EAHB SIG – Experimental Analysis of Human Behavior – Special Interest Group

FAPESP – Fundação de Amparo à Pesquisa do Estado de São Paulo

IRAP – Implicit Relational Assessment Procedure

JEAB – Journal of the Experimental Analysis of Behavior

JABA – Journal of Applied Behavior Analysis

LECH – Laboratório de Estudos sobre o Comportamento Humano

MTS – Matching-to-Sample

PPGpsi – Programa de Pós-Graduação em Psicologia

RFT – Relational Frame Theory (Teoria das Molduras Relacionais)

WEIRD – Western, Educated, Industrialized, Rich, and Democratic

JCBS – Journal of Contextual Behavioral Science

BAP – Behavior Analysis in Practice

BSI – Behavior and Social Issues

EOJBA – European Journal of Behavior Analysis

RCT – Randomized Controlled Trial

PRISMA – Preferred Reporting Items for Systematic Reviews and Meta-Analyses

### Apêndice III

#### Roteiro de Instruções Fornecidas às Crianças por Instrumento

Todas as instruções foram fornecidas individualmente, em ambiente escolar reservado, com apenas a pesquisadora e a criança presentes. Nenhuma informação sobre raça ou sobre os objetivos do estudo foi fornecida antes ou durante a aplicação dos instrumentos. Não foi fornecido feedback avaliativo sobre as respostas em nenhuma das medidas.

#### Teste Relacional (Matching-to-Sample AC3)

##### Instrução

Antes do início do teste, a pesquisadora verificou a compreensão dos símbolos A1 (polegar levantado, positivo) e A2 (polegar abaixado, negativo), conforme descrito nos estudos. A pesquisadora apresentava cada símbolo individualmente na tela e perguntava:

*Você sabe o que esse símbolo quer dizer?*

A conduta variava conforme a resposta da criança:

- Resposta correta: a pesquisadora reforçava — *Isso mesmo!*
- Resposta incompleta: a pesquisadora complementava sem corrigir diretamente
- Sem resposta ou resposta incorreta: a pesquisadora explicava — *Esse aqui é o positivo, que as pessoas usam quando gostam de alguma coisa, quando algo é bom (para A1) e esse aqui é o negativo, que as pessoas usam quando não gostam de alguma coisa, quando algo é ruim (para A2)*

Após confirmar a compreensão acerca do símbolo positivo e negativo, a experimentadora apresentou a seguinte instrução:

*Olha esse símbolo aqui no alto da tela. Você vai clicar nele. Depois vão aparecer três figuras embaixo. Você vai clicar na figura que você acha que combina mais com esse símbolo. Pode escolher qualquer uma das três figuras.*

### **Conduta durante as tentativas**

A pesquisadora permaneceu em silêncio durante as 16 tentativas. Não houve feedback após as respostas. Em caso de hesitação prolongada, a instrução reconstituída era:

*Pode escolher qualquer uma das três figuras que aparecem embaixo.*

### **Escala de Valência Emocional (Escala Likert de 5 pontos)**

A pesquisadora forneceu instruções verbais para orientar o uso da escala. A instrução reconstituída era:

*Vou mostrar para você algumas fotos de pessoas. Para cada foto, você vai me dizer o que sentiu ao ver aquela pessoa, usando essa escala aqui embaixo. Esse rostinho [aponta para o ponto 5, emoji muito feliz] quer dizer que você se sentiu muito bem ao ver essa pessoa. Esse [aponta para o 4] quer dizer que se sentiu bem. Esse do meio [aponta para o 3] quer dizer que não sentiu nem bem nem mal. Esse [aponta para o 2] quer dizer que se sentiu um pouco mal. E esse [aponta para o 1] quer dizer que se sentiu muito mal. Para cada foto, você vai escolher o rostinho que mostra o que você sentiu.*

### **Conduta durante as tentativas**

A pesquisadora permaneceu em silêncio. Não houve feedback sobre as respostas. A criança deslizava ou clicava sobre o ponto desejado.

### **Escala Self-Assessment Manikin (SAM)**

#### **Dimensão Valência**

*Você vai me dizer o quanto gostou ou não gostou dessa pessoa ao vê-la. Esse boneco aqui [aponta para o 5] tem um sorriso bem grande — quer dizer que você gostou muito. Esse [aponta para o 4] tem um sorriso menor — gostou um pouco. Esse do meio [aponta para o 3] não está sorrindo nem triste — não gostou nem desgostou. Esse [aponta*

*para o 2] está um pouco triste — desgostou um pouco. E esse [aponta para o 1] está muito triste — desgostou muito. Qual desses bonecos mostra como você se sentiu ao ver essa pessoa?*

### **Dimensão Ativação (Arousal)**

*Agora me diga se essa pessoa te deixa calmo ou te deixa com medo, nervoso, agitado. Esse boneco [aponta para o 5] está bem tranquilo, calmo, sem medo nenhum. Esse [aponta para o 4] está quase calmo. Esse do meio [aponta para o 3] não está nem calmo nem agitado. Esse [aponta para o 2] está um pouquinho assustado ou nervoso. E esse [aponta para o 1] está muito assustado, com muito medo. Qual desses bonecos mostra como essa pessoa te faz sentir?*

### **Dimensão Dominância**

*Agora me diga se essa pessoa parece forte e no controle, ou se parece fraca e sem controle. Esse boneco [aponta para o 5] é muito forte, parece que manda em tudo. Esse [aponta para o 4] é forte. Esse do meio [aponta para o 3] não é nem forte nem fraco. Esse [aponta para o 2] é um pouco fraco. E esse [aponta para o 1] é muito fraco, não manda em nada. Qual desses bonecos parece com essa pessoa da foto?*

### **Teste da Boneca — Versão Clássica Adaptada**

#### **Perguntas avaliativas — Estudo 4 (14 perguntas)**

Verificação da compreensão dos adjetivos positivos: boa, legal, inteligente, bonita, forte, confiável, superior.

*Você sabe o que significa [Adjetivo]?*

Verificação da compreensão dos adjetivos negativos: má, chato, pouco inteligente, feia, fraca, perigosa, inferior.

*Você sabe o que significa [Adjetivo]?*

*Finja que essas bonecas são pessoas. Vou te fazer algumas perguntas sobre elas. Para cada pergunta, você vai apontar para a boneca que você acha que combina com o que eu estou perguntando.*

Instrução reconstituída para cada adjetivo:

*Qual dessas pessoas é [adjetivo]?*

### **Perguntas de identificação racial**

*Qual dessas bonecas é branca? Qual é negra? Qual se parece com você?*

### **Conduta durante a aplicação**

A pesquisadora registrou manualmente no roteiro semiestruturado a boneca apontada (branca, negra, nenhuma, ou as duas) para cada adjetivo. Justificativas verbais espontâneas foram registradas literalmente. Não foi fornecido feedback sobre as escolhas.

### **Teste da Boneca Modificado (5 tonalidades de pele)**

*Olha essas bonecas aqui na tela. Vou te fazer algumas perguntas sobre elas. Imagine que essas bonecas são crianças novas que acabaram de chegar à sua escola. Para cada pergunta, você pode apontar para quantas quiser — pode ser nenhuma, uma, algumas ou todas. Não tem resposta certa ou errada.*

### **Perguntas — Estudo 3**

*Alguma você escolheria para ser sua amiga?*

*Alguma você escolheria para fazer um trabalho para nota*

*Alguma você escolheria para o time da Educação Física?*

*Alguma você escolheria para o seu grupo?*

### **Perguntas — Estudo 6**

*Alguma você escolheria para ser sua amiga?*

*Alguma você escolheria para fazer um trabalho para nota?*

*Alguma você escolheria para o time da Educação Física?*

*Alguma você desconfia que roubaria suas coisas?*

### **Conduta durante a aplicação**

Cada pergunta foi tratada como situação independente. A pesquisadora não comentou nem reagiu às escolhas. Em caso de hesitação, a instrução reconstituída era:

*Pode apontar para quantas quiser, ou para nenhuma.*

### **Function Acquisition Speed Test (FAST)**

A instrução a seguir foi padronizada com base no protocolo original de O'Reilly et al. (2012) e Cartwright et al. (2016), conforme descrito nos estudos desta tese:

*Na próxima etapa, sua tarefa será aprender qual botão pressionar quando uma imagem aparecer na tela. IMPORTANTE: durante esta fase, você deve pressionar apenas a tecla 'Z' ou a tecla 'M'. Por favor, localize essas teclas no teclado agora. Para ajudá-lo(a) a aprender, você receberá feedback informando se sua resposta está correta ou incorreta. Se tiver alguma dúvida, pergunte agora. Pressione qualquer tecla quando estiver pronto(a) para começar.*

### **Bloco de prática**

Antes dos blocos de teste, foi aplicado um bloco de prática com estímulos não relacionados ao estudo (ex.: partes do corpo, roupas, objetos), com o objetivo de familiarizar a criança com o funcionamento da tarefa sem introduzir os estímulos raciais ou avaliativos.

### **Estrutura dos blocos de teste**

**Bloco consistente:** tecla Z correta para faces negras e estímulos negativos; tecla M correta para faces brancas e estímulos positivos.

**Bloco inconsistente:** tecla Z correta para faces negras e estímulos positivos; tecla M correta para faces brancas e estímulos negativos.

### **Conduta durante a aplicação**

A pesquisadora permaneceu em silêncio e fora do campo visual da tela. Não foram fornecidos comentários sobre o desempenho. Em caso de distração ou interrupção, a instrução reconstituída era:

*Pode continuar. Pressione qualquer tecla para retomar.*

### **Encerramento de Sessão**

Ao final de cada sessão, independentemente dos instrumentos aplicados, a pesquisadora encerrava com agradecimento verbal padronizado. A instrução reconstituída era:

*Terminamos por hoje! Você nos ajudou muito. Obrigada por participar.*