

Revisão de estudos do Implicit Relational Assessment Procedure sobre vieses raciais¹

*(Review of the Implicit Relational Assessment Procedure
(IRAP) studies on racial biases)*

Táhcita Medrado Mizael*¹ & João Henrique de Almeida*

Universidade Federal de São Carlos

RESUMO

O estudo de vieses, estereótipos e preconceitos tem sido frequente nas pesquisas da psicologia. Uma preocupação frequente na área tem sido o nível de acurácia presente no relato dos participantes sobre seus preconceitos, uma vez que o autorrelato pode estar sujeito a vieses de deseabilidade social. Isso levou ao desenvolvimento de alternativas para a identificação desses vieses, por meio das chamadas medidas implícitas. O objetivo deste estudo foi revisar pesquisas empregando o Implicit Relational Assessment Procedure (IRAP) para investigar vieses, estereótipos e preconceitos raciais. Duas bases de dados (PubMed e Periódicos CAPES) e 18 periódicos (3 nacionais e 15 internacionais) foram consultados para a revisão. As palavras-chave utilizadas foram “Implicit Relational Assessment Procedure”, “Procedimento de Avaliação Relacional Implícita” e “IRAP”. Sete artigos foram encontrados. As pesquisas revisadas demonstraram a presença de vieses pró-branco e/ou anti-negro para os participantes brancos. Nas medidas explícitas, por outro lado, os participantes geralmente demonstraram vieses positivos para ambos os grupos. A coerência ou não entre as medidas implícitas e explícitas e a confiabilidade dos resultados foram discutidos utilizando-se modelos da Teoria das Molduras Relacionais. Em suma, o IRAP pode ser considerado um instrumento que, apesar de estar em desenvolvimento, possui resultados promissores que podem ser úteis e, inclusive, mais fidedignos que algumas medidas explícitas, especialmente nos contextos de pesquisa com temas socialmente sensíveis, como investigações sobre estereótipos e preconceitos.

1) Endereço para correspondência: Táhcita Medrado Mizael – Laboratório de Estudos sobre o Comportamento Humano (LECH) – Edifício Carolina Bori, Universidade Federal de São Carlos, Rodovia Washington Luís, km 235 - SP-310 - CEP 13565-905, São Carlos, São Paulo, Brasil. tahcitammizael@gmail.com

Táhcita M. Mizael é bolsista da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, Processo No. 15/10225-5). João Henrique de Almeida é bolsista da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, Processo No. 01874-7). Os autores são afiliados ao Instituto Nacional de Ciência e Tecnologia sobre Comportamento, Cognição e Ensino (INCT-ECCE), apoiado pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Processo No. 573972/2008-7) e pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP, Processo No. 2008/57705-8).

Palavras-Chave: Implicit Relational Assessment Procedure; IRAP; vieses raciais; preconceito racial; Procedimento de Avaliação Relacional Implícita.

ABSTRACT

Investigations of biases, stereotypes, and prejudices have been a frequent theme in psychology studies. A frequent concern in the field has been the level of accuracy present in the participant's reports on their prejudices. Taking into account the difficulties evidenced in self-report measures, which may be subject to social desirability biases, researchers have been developing alternative methodologies for identifying these biases, through the so-called implicit measures. In behavior analysis, some implicit measures were developed, such as the Implicit Relational Assessment Procedure (IRAP). This paper aimed to review studies that used the IRAP to investigate racial biases, stereotypes, and prejudices. Two databases (PubMed and Periódicos CAPES), in addition to 18 journals (three nationals and 15 internationals) were consulted. The keywords were "Implicit Relational Assessment Procedure" and "IRAP" and their respective terms in Portuguese. We found seven articles. From these, six used adults as participants. One study used children as participants. All studies investigated racial biases among Black and White individuals, except for one study that investigated biases of various ethnic groups (Irish, Scottish, American, and African). Regarding the results obtained in the IRAP, the reviewed studies demonstrated the presence of pro-White and/or anti-Black racial biases for White participants. On the other hand, on the explicit measures, participants generally showed positive biases for both groups. In sum, the IRAP can be considered an instrument that, despite being in development, has promising results that may be useful and even more reliable than some explicit measures, especially in socially sensitive research contexts such as investigations of stereotypes and prejudices.

Keywords: Implicit Relational Assessment Procedure; IRAP; racial bias; racial prejudice.

O estudo de vieses, estereótipos e preconceitos têm sido tema frequente nos estudos da Psicologia, sendo considerado, inclusive, como uma das temáticas principais de algumas áreas, como a Psicologia Social (e.g., Freedman, Carlsmith, & Sears, 1970; Schwarz & Bohner, 2001). Um dos preconceitos bastante pesquisados na Psicologia, em suas diversas áreas, tem sido o preconceito racial, que pode ser definido como uma atitude negativa com relação a indivíduos que possuem certas características físicas, como a cor da pele e outros marcadores sociais (e.g., avaliar negativamente certos formatos de nariz e/ou tipos de cabelo; Nogueira, 2006).

Apesar da existência de diversos estudos sobre a temática incluindo, mas não restrito, as teorias sobre formação e mudança de atitudes e estudos experimentais sobre o tema (e.g., Paluck & Green, 2009; Richeson & Sommers, 2015), uma preocupação frequente dos pesquisadores tem sido o nível de acurácia presente no relato dos participantes sobre seus preconceitos. Em relação ao uso do próprio relato dos participantes, Skinner (1953) já adver-

tia sobre essa prática, que apesar de ser amplamente utilizada tanto na Psicologia como em outras áreas das Ciências Sociais, pode representar um problema em relação a fidedignidade da informação. O autorrelato é um comportamento verbal (no contexto deste trabalho, entendido como uma resposta relacional derivada; Hayes, Barnes-Holmes, & Roche, 2001) que descreve o comportamento verbal ou não-verbal do próprio indivíduo, e este está sob controle de quaisquer variáveis ambientais presentes no momento de sua ocorrência (levando-se em consideração também a história de reforçamento do indivíduo). A precisão deste relato em relação aos padrões de respostas presentes no repertório comportamental de um indivíduo dependerá então das contingências em vigor, visto que este comportamento é regido pelos mesmos princípios comportamentais que qualquer outro e, dessa forma, as contingências de reforçamento e punição para o autorrelato devem ser observadas (Kohlsdorf & Costa Junior, 2009; Skinner, 1953). Uma vez que “ser preconceituoso” é geralmente conseqüenciado com reprovação, além de ser punível legalmente em alguns países (e.g., injúria racial no Brasil), é possível inferir que em grande parte dos casos a correspondência do autorrelato e do comportamento será baixa, graças às contingências. Esta baixa fidedignidade do autorrelato é um dos maiores motivadores para o desenvolvimento de metodologias alternativas para a identificação desses vieses, por meio das chamadas “medidas implícitas” (e.g., de Houwer, 2003; Fazio, Jackson, Dunton, & Williams, 1995; Greenwald, McGhee, & Schwartz, 1998; Nosek & Banaji, 2001).

A diferenciação das medidas implícitas e explícitas depende diretamente da perspectiva explicativa empregada. Em um escopo amplo (em que poderíamos incluir a Psicologia Social e a Psicologia Cognitiva), as medidas explícitas são medidas diretas que requerem que os participantes relatem seus julgamentos deliberadamente, enquanto as medidas implícitas captam avaliações automáticas, não declarativas e imediatas (de Houwer, 2006; Greenwald & Banaji, 1995). Além disso, poderíamos dizer também que as medidas explícitas são aquelas em que os participantes respondem livremente, sem restrições temporais, sobre seus preconceitos e vieses em questionários, escalas e outros instrumentos de autorrelato. De acordo com de Houwer (2006), as medidas podem ser consideradas implícitas quando a) os participantes não estão conscientes sobre a atitude que está sendo medida; b) não têm acesso consciente/deliberado/intencional à atitude (*conscious access*); e/ou c) não podem controlar os resultados obtidos pela medida. Estas diferenciações possuem detalhes muito mais intrincados a depender se a explicação é associacionista ou proposicionista, pois descrevem de forma diferente como esses comportamentos são estabelecidos (para maiores informações, ver Hughes, Barnes-Holmes, & de Houwer, 2011).

No escopo do trabalho atual a diferenciação relevante se dá não pelo controle do comportamento (se deliberado ou involuntário), uma vez que, em uma perspectiva comportamental, seguiremos a proposta skinneriana de que o comportamento é determinado pelo seu ambiente (Skinner, 1953). A diferença entre elas também não é relativa ao tipo de comportamento descrito, uma vez que este comportamento descrito como “atitude” será visto aqui como um comportamento verbal como definido pela Teoria das Molduras Relacionais, ou seja, uma Resposta Relacional Arbitrariamente Aplicável² (RRAA; Hayes et al., 2001). A

2) O Responder Relacional Arbitrariamente Aplicável é a unidade comportamental que permite fazer uma análise puramente funcional do comportamento verbal, na perspectiva da Teoria das Molduras Relacionais. Este comportamento é derivado, controlado por dicas contextuais, e frequentemente exibe uma transformação de funções coerente com a rede relacional estabelecida (para mais informações, ver Hayes et al., 2001).

diferenciação entre explícito ou implícito, neste caso, está em quão rápido estas respostas relacionais são evocadas. Vamos imaginar que uma pessoa encontre um estímulo que tenha alguma função social carregada (*socially loaded*). Respostas relacionais irão ocorrer de forma relativamente rápida ou mais demorada. Da perspectiva do modelo de Elaboração e Coerência Relacional (REC; *Relational Elaboration and Coherence*), as respostas relacionais breves e imediatas (BIRRs, *Brief and Immediate Relational Responding* em inglês) seriam a base do que os pesquisadores chamam de atitudes implícitas. Por outro lado, as respostas elaboradas e estendidas (EERRs, *Extended and Elaborate Relational Responding* em inglês) que forem coerentes com uma ou mais respostas relacionais no repertório comportamental desta pessoa representaria o que é chamado de atitudes explícitas. A diferença se dá somente nestes dois padrões amplos de respostas relacionais frequentemente observados (Barnes-Holmes, Barnes-Holmes, Stewart, & Boles, 2010; Hughes et al., 2011).

Na Análise do Comportamento, algumas medidas implícitas foram desenvolvidas, como o Teste Rápido de Velocidade de Aquisição (FAST – *Fast Acquisition Speed Test*; O’Reilly, Roche, Ruiz, Tyndall, & Gavin, 2012) e a Tarefa de Responder Relacional (RRT – *Relational Responding Task*; de Houwer, Heider, Spruyt, Roets, & Hughes, 2015). Uma medida “implícita” proposta por analistas do comportamento que tem tido destaque em várias publicações é o Procedimento de Avaliação Relacional Implícita (*Implicit Relational Assessment Procedure*, IRAP – Barnes-Holmes, et al., 2006).

O IRAP é um instrumento que apresenta termos relacionais aos participantes, exigindo que se responda de maneira rápida e acurada a relações consideradas consistentes e inconsistentes com a história de vida dos participantes. Baseado na hipótese de que se responde mais rápido à relações consistentes do que à relações inconsistentes com nossa história de vida, a latência de resposta é central para a identificação dos vieses dos participantes (Barnes-Holmes, et al., 2006; Power, Barnes-Holmes, & Barnes-Holmes, & Stewart, 2009).

No IRAP, em cada tentativa, um estímulo-alvo (conjuntos ou pares de estímulos que serão avaliados um a um) é apresentado juntamente com rótulos (elementos que descrevem propriedades significativas; na maioria das vezes adjetivos, mas em alguns casos podem ser outros estímulos que possuem alguma valência). Esses elementos devem ser relacionados a cada tentativa e o controle contextual relacional e funcional evocará uma resposta relacional breve e imediata, uma vez que o participante tem, geralmente, dois segundos em média para responder a cada relação. Também são apresentadas duas opções de resposta (e.g., verdadeiro/falso; similar/oposto; sim/não), utilizadas para que o participante confirme ou negue a relação entre rótulo e estímulo-alvo.

É importante ressaltar que alvos e rótulos podem ter suas posições alteradas. Em muitos experimentos em países de língua inglesa o rótulo (normalmente um adjetivo) vem na parte superior da tela e o estímulo-alvo (objeto da análise) no centro, como é frequente neste idioma. No Brasil, sabe-se do uso do IRAP mantendo essa forma de apresentação dos estímulos, como também em uma apresentação culturalmente mais lógica e coerente com nosso idioma, i.e., com o estímulo-alvo na parte superior e o rótulo na parte central da tela do computador. Essas diferenças não são um problema, uma vez que não afetam propriedades funcionais das respostas relacionais evocadas. Além disso, é importante explicitar que as opções de resposta (1 e 2 na Figura 1) constantemente alternam os cantos inferiores da tela em quase toda a literatura do IRAP, com poucos estudos no qual a ausência dessa variação foi investigada (ver, por exemplo, Campbell, Barnes-Holmes, Barnes-Holmes, & Stewart, 2011). Na Figura 1 é

possível observar a disposição dos estímulos na tela no padrão tradicional (quadro à esquerda) e também um exemplo de uma tentativa (quadro à direita).



Figura 1. Estrutura tradicional de uma tentativa no Procedimento de Avaliação Relacional Implícita (IRAP) e exemplo de uma tentativa. As posições de “rótulo” e “alvo” em alguns procedimentos são invertidas, como comentado no texto. Com frequência, as opções de resposta alternam posições (se direita ou esquerda) aleatoriamente a cada tentativa.

As tentativas são separadas em blocos, sendo alguns considerados consistentes com a história de vida dos participantes, e outros blocos inconsistentes com tal história. Os participantes devem responder de acordo com uma regra dada pelo experimentador, que é alternada a cada bloco. Em cada tentativa, os acertos são seguidos pela apresentação da próxima tentativa, enquanto erros são consequenciados com um X vermelho no centro da tela, de modo que a próxima tentativa só aparece quando os participantes respondem de acordo com a relação estabelecida pelo experimentador naquele bloco específico. Geralmente, os participantes são submetidos a, no mínimo, dois blocos de treino, antes de serem encaminhados para a fase de testes, que geralmente conta com seis blocos. A fase de testes costuma ser idêntica à fase de treino, ou seja, blocos consistentes e inconsistentes são apresentados de maneira alternada. Os critérios utilizados nos blocos de treino podem variar, mas em grande parte dos estudos conduzidos com adultos, tem-se requerido, no mínimo, 80% de acertos em cada bloco, com uma latência média de respostas de até dois segundos em cada tentativa, sendo necessário atingir esse critério em dois blocos consecutivos (i.e., um consistente e um inconsistente, ou vice-versa, para ir para a fase de testes; Vahey, Nicholson, & Barnes-Holmes, 2015).

O IRAP tem sido utilizado para avaliar uma variedade de vieses e preconceitos, como estereótipos de gênero (e.g., Cartwright, Hussey, Roche, Dunne, & Murphy, 2016; Farrel & McHugh, 2017; Rabelo, Bortoloti, & Souza, 2014), vieses raciais (e.g., Barnes-Holmes, Barnes-Holmes, & Stewart, 2010; Power, Harte, Barnes-Holmes e Barnes-Holmes, 2017a), gordofobia (e.g., Drake et al., 2010; Ritzert et al., 2016), entre outros. Levando em consideração a relevância não apenas científica, mas também social do estudo de vieses e preconceitos raciais, além da importância de se desenvolver e utilizar as chamadas medidas implícitas (medidas obtidas a partir de BIRRs), o objetivo deste trabalho foi revisar os estudos que utilizaram o IRAP para investigar vieses, estereótipos e preconceitos raciais.

MÉTODOS

Duas bases de dados (PubMed e Periódicos CAPES), além de 18 periódicos, três nacionais e 15 internacionais foram consultados para a revisão. Os periódicos nacionais foram: Revista Brasileira de Análise do Comportamento (REBAC), Revista Brasileira de Terapia Comportamental e Cognitiva (RBTCC), e Revista Perspectivas em Análise do Comportamento; e os internacionais foram: Acta Comportamental, Behavior Analysis in Practice, Behavior and Social Issues (BSI), Behavior Analysis: Research and Practice, Behavioral Interventions, Behavioral Processes, European Journal of Behavior Analysis (EJOBA), Journal of Applied Behavior Analysis (JABA), Journal of Contextual Behavioral Science, Journal of the Experimental Analysis of Behavior (JEAB), Learning and Behavior, Psychology and Neuroscience, The Analysis of Verbal Behavior, The Behavior Analyst, e The Psychological Record. Esses periódicos foram selecionados por serem periódicos que publicam trabalhos da área (Análise do Comportamento).

As palavras-chave buscadas foram: “*Implicit Relational Assessment Procedure*” e “IRAP”. Caso a busca resultasse em mais de 50 artigos, as palavras-chave “*race*” e “*racial*” eram adicionadas. Nos periódicos brasileiros, além destas palavras, também foram buscadas “Procedimento de Avaliação Relacional Implícita” e “*implícita*” ou “*implícito*”. Todos os artigos que utilizaram o IRAP para investigar questões raciais (e.g., vieses, estereótipos, atitudes), independente da data de publicação, foram incluídos. A busca foi realizada em dezembro de 2017.

RESULTADOS

Sete artigos foram encontrados. Destes, seis foram publicados no periódico *The Psychological Record* (Power et al. 2009; Barnes-Holmes et al., 2010; Drake et al., 2010; Mizael, de Almeida, Silveira, & de Rose, 2016; Power et al., 2017a; Power, Harte, Barnes-Holmes, & Barnes-Holmes, 2017b) e um no *Behavior and Social Issues* (Drake et al., 2015).

Dos sete estudos encontrados, seis utilizaram como participantes adultos, geralmente estudantes universitários. Apenas um estudo utilizou, como participantes, crianças (Mizael et al., 2016). Todos os estudos investigaram vieses raciais entre indivíduos brancos e negros, com exceção da pesquisa de Power et al. (2009), que investigou vieses de diversos grupos étnicos (irlandeses, escoceses, americanos e africanos). Considerando os periódicos que publicaram estes trabalhos, além de serem publicados majoritariamente no mesmo periódico, *The Psychological Record*, os mesmos pesquisadores têm sido autores de várias publicações. Por exemplo, Drake, Harte, Murphy e Stewart foram autores de duas publicações; Power, de três, e Barnes-Holmes, D. e Barnes-Holmes, Y. de quatro das sete publicações, o que mostra que tais estudos têm sido uma preocupação de grupos de pesquisa ainda restritos.

A Tabela 1 resume os artigos encontrados, enumerando os autores, objetivo, número e idade dos participantes recrutados, tipo de medida obtida, estímulos utilizados no IRAP, se houve consistência entre as medidas implícitas e explícitas (quando ambas foram empregadas) e principais resultados.

A seguir, encontra-se um resumo de cada trabalho encontrado.

Tabela 1. Resumo dos estudos encontrados

Autores	Objetivo	N	Idade	Tipo de Medida Obtida	Estímulos utilizados no IRAP	Critérios de Latência (L) e Acurácia (A):	Resumo dos Resultados
Power, Barnes- Holmes, Barnes- Holmes e Stewart (2009)	Verificar se o IRAP poderia ser usado na identificação de atitudes socialmente sensíveis, além de determinar se o IRAP poderia ser utilizado para avaliar relações comparativas.	32 (16 no Estudo 1 e 16 no Estudo 2)	18-35	Escala Likert e IRAP	Estímulos-Alvo: irlandeses/ escoceses/ americanos/ africanos Rótulos: mais agradável-menos agradável Opções de Resposta: verdadeiro/falso	L ≤ 3s A = NI Não	Estudo 1 - IRAP: Preferência por irlandeses sobre escoceses, escoceses sobre americanos, e americanos sobre africanos. Escala Likert: Igual preferência por irlandeses e escoceses, e uma preferência maior por africanos do que por americanos. Estudo 2 - IRAP: Preferência por americanos sobre irlandeses, irlandeses sobre escoceses, e escoceses sobre africanos. Escala Likert: Preferência por irlandeses sobre todos os outros grupos, e uma preferência similar por americanos, escoceses e africanos.

Autores	Objetivo	N	Idade	Tipo de Medida Obtida	Estímulos utilizados no IRAP	Critérios de Latência (L) e Acurácia (A):	Consistência MI e ME	Resumo dos Resultados
Barnes-Holmes, Murphy, Barnes-Holmes e Stewart (2010)	Avaliar estereótipos implícitos pró-branco e anti-negro em participantes irlandeses em um contexto público ou privado.	50 (31 no Estudo 1 e 19 no Estudo 2)	18-62 (Estudo 1) 18-52 (Estudo 2)	Escala de discriminação (DS) e diversidade (DV), escala de racismo moderno modificada (MMRS), escalas Likert e IRAP	Estímulos-Alvo: negro/branco Rótulos: seguro- perigoso Opções de Resposta: verdadeiro/falso	Estudo 1: L ≤ 3s A ≥ 80% Estudo 2: L ≤ 2s A ≥ 80%	Não	Estudo 1 – IRAP: O contexto público diminuiu o uso de estereótipos pró-branco nas tentativas “negro- seguro” em comparação com o contexto privado, mas isso não ocorreu para todas as tentativas. Medidas Explícitas: Vieses positivos com relação aos negros para ambos os grupos ou vieses pró-negro maiores na condição público. Estudo 2 – IRAP: A redução da latência de respostas aumentou significativamente o viés pró-branco nas tentativas “branco-seguro”, além de produzir vieses anti- negro significativos nas tentativas “negro- perigoso”, as quais não foram observadas no Estudo 1. Medidas Explícitas: Mesmo dado obtido no Estudo 1.

Autores	Objetivo	N	Idade	Tipo de Medida Obtida	Estímulos utilizados no IRAP	Critérios de Latência (L) e Acurácia (A):	Consistência MI e ME	Resumo dos Resultados
Drake, Kellum, Wilson, Luoma, Weinstein e Adams (2010)	Examinar a sensibilidade e aplicabilidade do IRAP em quatro temas: raça, religião, gênero e obesidade.	67 (15 na condição "raça")	Idade média = 19 anos de idade	IRAP	Estímulos-Alvo: branco/negro Rótulos: positivo-negativo Opções de Resposta: verdadeiro/falso	L= NI A ≥ 65%	Não avaliada	Nas tentativas "branco-positivo", os participantes responderam significativamente mais rápido com "verdadeiro" do que com "falso". Nas tentativas "negro-negativo", os resultados não foram significativos, assim como nas tentativas "negro-positivo".

Autores	Objetivo	N	Idade	Tipo de Medida Obtida	Estímulos utilizados no IRAP	Critérios de Latência (L) e Acurácia (A):	Consistência MI e ME	Resumo dos Resultados
Drake Kramer, Sain, Swiatek, Kohn e Murphy (2015)	Explorar a confiabilidade e a validade convergente do IRAP com uma amostra balanceada de pessoas brancas e negras	57 (22 negros, 22 brancos e 13 com classificação racial mista ou distinta	Idade média = 19 anos de idade	Escala de racismo moderno (MRS), escala de orientação de dominação social (SDO), diferencial semântico e IRAP	Estímulos-Alvo: branco/negro Rótulos: positivo-negativo Opções de Resposta: verdadeiro/falso	L ≤ 2s A ≥ 80%	Não identificada	Na primeira aplicação do IRAP, os grupos branco e negro mostraram vieses a favor do próprio grupo, sendo os vieses dos brancos em maior grau que dos negros. No geral, esses dois grupos mostraram pouco ou nenhum viés positivo ou negativo com relação ao outro grupo. Na segunda aplicação, o "grau" dos vieses a favor do próprio grupo diminuiu para todos os grupos. Além disso, os negros mostraram viés positivo para os brancos e os brancos, viés positivo para os negros. O grupo com classificação racial mista ou distinta teve resultados mistos. Nas medidas explícitas, os negros demonstraram o menor nível de racismo sutil e de dominância social e os brancos, o maior nível.

Autores	Objetivo	N	Idade	Tipo de Medida Obtida	Estímulos utilizados no IRAP	Critérios de Latência (L) e Acurácia (A):	Consistência MI e ME	Resumo dos Resultados
Mizael, de Almeida, Silveira e de Rose (2016)	Investigar se uma otimização em procedimentos de treino para reverter classes de equivalência com estímulos socialmente sensíveis aumentaria a probabilidade de os participantes formarem classes de equivalência entre faces de pessoas negras e um símbolo considerado positivo.	13	8-10	Auto-Avaliação Manikin (SAM) e IRAP	Estímulos-Alvo: negro/branco Rótulos: figuras positivas-figuras negativas Opções de Resposta: sim/não	L ≤ 3s A ≥ 70%	Sim	Todas as crianças formaram as classes de equivalência, passando a relacionar as faces negras com o símbolo positivo. Os resultados do IRAP mostraram ausência de vieses raciais negativos com relação a pessoas brancas e negras no pós-teste, confirmando os resultados obtidos no SAM no pós-teste, no qual a diferença entre as avaliações de faces brancas e negras não foi estatisticamente significativa.

Autores	Objetivo	N	Idade	Tipo de Medida Obtida	Estímulos utilizados no IRAP	Critérios de Latência (L) e Acurácia (A):	Consistência MI e ME	Resumo dos Resultados
Power, Harte, Barnes- Holmes e Barnes- Holmes (2017a)	Analisar vieses raciais de brancos e negros residentes da Irlanda.	34 (16 negros e 18 brancos)	17-28	Escala de discriminação (DS) e diversidade (DV), escala de diferencial semântico, “ <i>feeling thermometers</i> ” e IRAP	Estímulos-Alvo: “Acho que pessoas negras/brancas são...” Rotulos: atributos positivos-atributos negativos Opções de Resposta: verdadeiro/falso	L ≤ 2s A ≥ 80%	Sim, para participantes negros. Parcial para participantes brancos.	IRAP: Os participantes negros demonstraram vieses pró-branco e pró- negro em todos os quatro tipos de tentativas; os participantes brancos demonstraram vieses positivos nos dois tipos de tentativas que envolviam o rótulo “pessoas brancas”, além das tentativas do tipo “negro-positivo”, mas mostraram também um forte viés negativo nas tentativas do tipo “negro- negativo”. Medidas Explícitas: No geral, as escalas mostraram vieses pró- negro tanto para os participantes negros quanto brancos, com vieses positivos para o próprio grupo maiores que para o outro grupo.

Autores	Objetivo	N	Idade	Tipo de Medida Obtida	Estímulos utilizados no IRAP	Critérios de Latência (L) e Acurácia (A):	Consistência MI e ME	Resumo dos Resultados
Power, Harte, Barnes-Holmes e Barnes-Holmes (2017b)	Avaliar vieses raciais de irlandeses brancos no IRAP e comparar com medidas de eletroencefalograma (EEG) obtidas enquanto os participantes faziam o IRAP.	16 (todos brancos)	18-33	IRAP e EEG	Estímulos-Alvo: "Eu acho que pessoas NEGRAS/BRANCAS são" Rótulos: atributos positivos-atributos negativos Opções de Resposta: verdadeiro/falso	L ≤ 2s A ≥ 80%	Não avaliada	IRAP: os participantes demonstraram vieses pró-branco e anti-negro. EEG: obtenção de padrões de atividade diferentes quando as tentativas eram consistentes, em comparação com as inconsistentes.

Nota. N= número de participantes. MI= medida implícita. ME= medida explícita. NI= não-identificado.

O estudo de Power et al. (2009) teve como objetivos verificar se o IRAP poderia ser usado na identificação de atitudes socialmente sensíveis, isto é, atitudes relacionadas a preconceitos, além de determinar se o IRAP poderia ser utilizado para avaliar relações de comparação (relações em que dois ou mais estímulos são comparados com relação a um ou mais atributos). Dezesesseis participantes irlandeses, de 18 a 35 anos de idade foram recrutados para o Estudo 1, sendo que foram investigadas como se davam relações de maior ou menor preferência com relação a grupos sociais na história de vida destes sujeitos. Em cada tentativa, as frases “mais agradável” (*more likeable*) ou “menos agradável” (*less likeable*) apareceriam como rótulos, e o alvo era um par de palavras dentre três pares (irlandês-escocês; escocês-americano; americano-africano). Em cada tentativa, os participantes tinham que responder se a relação apresentada era “verdadeira” ou “falsa”, de acordo com a regra pré-estabelecida, em até três segundos.

As relações consideradas consistentes com a história de vida dos participantes foram uma preferência por irlandeses sobre os escoceses, por escoceses sobre americanos, e por americanos sobre africanos. Essas relações foram baseadas em pesquisas que mostravam que indivíduos tendem a preferir seu próprio grupo do que outros (Tajfel, 1982), e que o grau de preferência tem relação com o quanto determinado grupo é percebido como similar a seu próprio (e.g., Goethals & Darley, 1977).

Antes de fazer o IRAP, porém, os participantes tinham que avaliar a preferência dos grupos sociais supracitados, em quatro escalas Likert com 13 pontos, que variavam de -6 (extremamente desagradável) a +6 (extremamente agradável). Os resultados do IRAP mostraram que os participantes demonstraram preferências por irlandeses sobre escoceses, escoceses sobre americanos, e americanos sobre africanos, exatamente como foi hipotetizado. Diferentemente, nas escalas Likert, os participantes declararam uma mesma preferência por irlandeses e escoceses, e uma preferência maior por africanos que por americanos.

No Estudo 2, o procedimento foi repetido com 16 cidadãos dos Estados Unidos, auto identificados como “irlandeses americanos”, com idades entre 18 e 35 anos. A única diferença, além dos participantes recrutados, foi os estímulos-alvo utilizados no IRAP. Dessa vez, os pares foram “americano-irlandês, irlandês-escocês e escocês-africano”. Essa ordem dos pares foi baseada em estudo piloto que indicou que irlandeses americanos geralmente avaliavam irlandeses como mais similares aos americanos, seguido pelos escoceses e africanos, nessa ordem.

De acordo com a hipótese apresentada, os resultados do IRAP mostraram que os participantes demonstraram uma preferência por americanos sobre irlandeses, irlandeses sobre escoceses, e escoceses sobre africanos. Novamente, na medida explícita os resultados divergiram do IRAP, indicando uma preferência por irlandeses sobre todos os outros grupos, e uma preferência similar por americanos, escoceses e africanos. Essa divergência entre as medidas explícitas e implícitas encontrada no Estudo 1 foi, portanto, replicada no Estudo 2, sendo considerada comum em pesquisas cujos estímulos utilizados são relacionados a um tópico socialmente sensível como o preconceito.

A pesquisa de Barnes-Holmes et al. (2010) teve a finalidade de avaliar estereótipos implícitos pró-branco e anti-negro em participantes irlandeses em um contexto público ou privado. Esse tipo de variável de contexto foi considerado relevante em estudos prévios com o Teste de Associação Implícita (IAT³, uma medida considerada implícita - e.g., Boysen, Vogel,

3) Em resumo, o IAT, instrumento proposto por Greenwald, McGhee e Schwartz em 1998, em um contexto de pesquisa na área da Psicologia Cognitiva, busca verificar a robustez de associações automáticas entre represen-

& Madon, 2006), no qual foi identificado um menor viés negativo sobre a homossexualidade no contexto público, em comparação com o contexto privado. No Estudo 1, 31 participantes irlandeses, com idades entre 18 e 62 anos, foram divididos em dois grupos: público (N=16) e privado (N=15). Todos os participantes sabiam que a pesquisa era sobre estereótipos raciais.

Três medidas de autorrelato também foram utilizadas: as escalas de discriminação (DS) e diversidade (DV), a escala de racismo moderno modificada (MMRS), e escalas Likert (para indicar o quão seguros ou perigosos homens brancos e negros eram). No IRAP, os rótulos utilizados foram “seguro” e “perigoso”, os estímulos-alvo eram seis fotos de homens segurando armas, três brancos e três negros, e as opções de resposta eram “verdadeiro” e “falso”. As relações consideradas consistentes com a história de vida dos participantes eram “seguro-branco-verdadeiro”, “perigoso-negro-verdadeiro”, “seguro-negro-falso” e “perigoso-branco-falso”. Os critérios utilizados foram uma latência média de até três segundos para cada resposta, com uma acurácia média de 80% de acertos em cada bloco.

Na condição “público”, era evidenciado aos participantes que eles fariam tarefas que indicariam o nível de viés que eles tinham com relação a pessoas negras. Em seguida, um dos experimentadores sentava ao lado do participante enquanto este fazia o IRAP, e por fim, os participantes tinham que verbalizar suas respostas nas medidas explícitas para que o experimentador as anotasse. Na condição “privado”, por outro lado, o experimentador ressaltava que iria coletar os dados do participante, mas não iria examiná-los. Os participantes faziam o IRAP sozinhos, e respondiam as escalas anotando eles mesmos suas crenças nas folhas.

Os resultados do IRAP mostraram que, de maneira geral, os participantes responderam “verdadeiro” mais rápido que “falso” nas tentativas “branco-seguro” e “negro-perigoso”, e “falso” mais rápido que “verdadeiro” nas tentativas “branco-perigoso”, o que confirma as hipóteses. Porém, nas tentativas “negro-seguro”, os participantes responderam “verdadeiro” mais rápido que quando responderam “falso”, contrariando tal hipótese. Houve diferenças nos contextos público e privado: o contexto público diminuiu o uso de estereótipos pró-branco nas tentativas “negro-seguro”, em comparação com o contexto privado.

Nas medidas explícitas, o MMRS indicou vieses positivos com relação aos negros para ambos os grupos. Nas escalas DS e DV, os resultados indicaram maiores vieses raciais positivos para negros na condição público do que privado, e nas escalas Likert, novamente, os resultados mostraram maior viés positivo para negros na condição público, em comparação com a condição privado. Não foram apresentados pelos pesquisadores os dados das medidas explícitas em relação a avaliação dos estímulos que apresentavam homens brancos com armas de fogo. Uma vez que os resultados do IRAP nos quatro tipos de tentativas não foram consistentes com a previsão realizada previamente a coleta, os autores hipotetizaram que uma maior latência para responder poderia ter influenciado os resultados, e de fato, os participantes do grupo “privado” tiveram uma média de latência de respostas maior que os do grupo “público”. Desse modo, o Estudo 2 teve como objetivo explorar os efeitos da restrição temporal nos resultados do IRAP.

Dezenove participantes irlandeses, com idades entre 18 e 52 anos realizaram novamente o mesmo procedimento que os participantes da condição público previamente mencionada. A única diferença era que a latência permitida no IRAP foi reduzida para, no máximo, dois

tações de conceitos na memória. Durante a tarefa, o participante é requerido a categorizar os estímulos apresentados na tela pressionando uma de duas teclas no computador. A latência é a medida básica de interesse, apesar de não haver restrição no tempo de resposta. Para maiores informações, ver Greenwald, Nosek e Banaji (2003).

segundos. Os resultados mostraram que a redução da latência de respostas aumentou significativamente o viés pró-branco nas tentativas “branco-seguro”, além de produzir vieses anti-negro significativos nas tentativas “negro-perigoso”, as quais não foram observadas no grupo com a latência de respostas de até três segundos. Portanto, a previsão de que aumentar a automaticidade, ao reduzir a latência de respostas, iria produzir maiores evidências de vieses raciais no IRAP foi confirmada.

A pesquisa de Drake et al. (2010) teve como propósito examinar a sensibilidade e aplicabilidade do IRAP em quatro temas: raça, religião, gênero e obesidade. Dada a temática do artigo, serão expostos somente os dados relativos ao IRAP de raça. Quinze universitários (dos 67 participantes recrutados) fizeram o IRAP sobre raça. Os estímulos alvo foram “branco” e “negro”, os rótulos positivos “bom”, “inteligente”, “superior”, “merecedor”, “motivado” e “digno”, e os negativos, “ruim”, “burro”, “inferior”, “inadequado”, “preguiçoso” e “deficiente” (*good, smart, superior, deserving, motivated, worthy, bad, stupid, inferior, inadequate, lazy e deficient*). As opções de resposta foram “verdadeiro” e “falso”. O critério utilizado foi uma acurácia de, no mínimo, 65% de acertos em cada bloco. Antes de realizar as tarefas, os participantes eram avisados que o feedback dado pelo computador poderia fazer sentido algumas vezes, outras não, mas que eles deveriam responder rapidamente e com o menor número de erros possível. Os resultados mostraram que, nas tentativas “branco-positivo”, os participantes responderam significativamente mais rápido com “verdadeiro” do que com “falso”. Com relação às tentativas envolvendo “negro” (negro-positivo e negro-negativo), estas não foram estatisticamente diferentes de zero, o que não permitiu afirmar a existência de um viés positivo ou negativo para os negros.

O trabalho de Drake et al. (2015) teve como objetivo explorar a confiabilidade e a validade convergente do IRAP com uma amostra balanceada de pessoas brancas e negras. Cinquenta e sete participantes foram divididos em três grupos: participantes negros (N= 22), brancos (N=22) e outros (mais de uma raça ou raça diferente destas; N=13). Foram sumarizados dados demográficos dos participantes (idade, raça, sexo e religião), além de serem utilizadas três medidas explícitas: a escala de racismo moderno (MRS), a escala de orientação de dominação social (SDO - analisava a cognição dos participantes sobre iniquidade entre grupos, não sendo específica de nenhuma raça) e o diferencial semântico para avaliar as palavras utilizadas no IRAP (e.g., positivo, negativo etc.).

Os participantes faziam o IRAP três vezes: a primeira era uma espécie de “pré-treino”, com palavras “neutras” (e.g., flor e esgoto, bom e ruim etc.), e as demais eram sobre raça. Os rótulos eram “pessoas negras” e “pessoas brancas”, os estímulos positivos “merecedor”, “bom”, “motivado”, “inteligente”, “superior” e “digno”, e negativos “ruim”, “deficiente”, “inadequado”, “inferior”, “preguiçoso” e “estúpido/burro” (*deserving, good, motivated, smart, superior, worthy, bad, deficient, inadequate, inferior, lazy e stupid*). As opções de resposta eram “verdadeiro” e “falso”, e o critério utilizado para contabilizar os resultados foi de, no mínimo, 80% de acertos em um bloco e dois segundos de latência média de respostas em cada tentativa. O primeiro IRAP de raça tinha dois blocos de treino, e o segundo era constituído somente pelos blocos de teste. A ordem de apresentação das tarefas (medidas demográficas + medidas explícitas e IRAP) foi contrabalanceada entre os participantes.

Com relação ao IRAP, somente 36 participantes atingiram os critérios nas duas aplicações do instrumento aplicadas à raça (nove negros, 18 brancos, e nove com outra classificação racial). De maneira resumida, na primeira aplicação do IRAP, os grupos branco e negro mostraram vieses a favor do próprio grupo, sendo os vieses dos brancos em maior

grau que dos negros, nas comparações dos valores dos escores D-IRAP. No geral, esses dois grupos mostraram pouco ou nenhum viés positivo ou negativo com relação ao outro grupo. Na segunda aplicação, o “grau” dos vieses a favor do próprio grupo diminuiu para todos os participantes brancos e negros. Além disso, os participantes negros mostraram viés positivo para os brancos e os participantes brancos, viés positivo para os negros. O grupo “outros” teve resultados mistos.

Nas medidas explícitas, os negros demonstraram o menor nível de racismo sutil e de dominância social, e os brancos o maior nível nos dois instrumentos. Os autores salientaram, na discussão, as vantagens do IRAP com relação ao IAT, mostrando que os vieses identificados na pesquisa (que se mostraram a favor do próprio grupo, mas não contra o outro grupo) não seriam identificados com o uso do IAT. Além disso, a mudança de viés identificada da primeira para a segunda aplicação do instrumento, na qual atitudes mais igualitárias começaram a aparecer, foi ressaltada, o que pode levantar questões sobre a estabilidade e a confiabilidade do IRAP, caso sejam feitas aplicações sucessivas do instrumento.

A pesquisa de Mizael et al. (2016) investigou se uma otimização em procedimentos de treino para reverter classes de equivalência com estímulos socialmente sensíveis aumentaria a probabilidade de participantes que demonstraram viés racial formarem classes de equivalência entre faces negras e um símbolo considerado positivo. Treze crianças, com idades entre oito e 10 anos e que demonstraram vieses raciais (a partir de um instrumento de autorrelato) foram recrutadas para a pesquisa.

O procedimento envolvia o treino de relações entre um símbolo positivo e um abstrato, e entre o abstrato e quatro faces de pessoas negras, em uma das classes, e entre um símbolo negativo e um estímulo abstrato diferente, e entre este e um terceiro estímulo abstrato, também distinto dos dois primeiros, na segunda classe. Os testes verificaram, portanto, se as crianças passariam a relacionar as faces de pessoas negras com o símbolo positivo, e o símbolo negativo ao abstrato correspondente. Todas as crianças formaram as classes de equivalência, passando a relacionar as faces negras com o símbolo positivo.

Após formarem as classes, o instrumento utilizado no pré-teste foi repetido e, em seguida, as crianças fizeram o IRAP. Os estímulos alvo foram quatro faces de pessoas negras e quatro faces de pessoas brancas (uma por vez e as mesmas utilizadas na fase de treino de formação de classes), os rótulos, quatro figuras consideradas positivas (criança segurando sorvete, dois golfinhos, barra de chocolate e filhotes de cachorro) ou negativas (frutas estragadas, menino sentindo dor, serpente e cachorro bravo). Os critérios utilizados foram de, no mínimo, 70% de acertos em cada bloco, e uma latência média de resposta de até três segundos. Esses critérios, menos rigorosos, se basearam em uma pesquisa prévia que utilizou o IRAP com crianças (e.g., Rabelo et al., 2014).

Os resultados do IRAP mostraram ausência de vieses raciais negativos com relação a pessoas brancas e negras no pós-teste, confirmando os resultados obtidos na medida explícita (autorrelato), no qual a diferença entre as avaliações de faces brancas e negras não foi estatisticamente significativa no pós-teste (no pré-teste, as crianças foram recrutadas justamente por demonstrarem diferenças estatisticamente significativas entre as avaliações das faces, com as faces brancas sendo mais bem avaliadas que as negras). Apesar de os resultados serem convergentes com a medida explícita, um dado que não costuma ocorrer quando se trata de tópicos socialmente sensíveis, os participantes foram crianças (essa divergência foi identificada em estudos com adultos), e o IRAP não foi realizado no pré-teste, impossibilitando saber, de fato, quais foram as diferenças obtidas neste instrumento antes e depois da formação das classes.

A pesquisa de Power et al. (2017a) teve a finalidade de analisar vieses raciais de indivíduos brancos e negros residentes da Irlanda. Trinta e quatro participantes, 16 negros e 18 brancos, com idades entre 17 e 28 anos foram recrutados para a pesquisa. Inicialmente, os participantes faziam o IRAP. As frases “Eu acho que pessoas brancas são” e “Eu acho que pessoas negras são” funcionaram como estímulos-alvo. Os rótulos utilizados foram “amigável”, “honesto”, “trabalhador”, “pacífico”, “bom” e “inteligente; “hostil”, “enganador/falso”, “preguiçoso”, “violento”, “ruim” e “burro” (*friendly, honest, hardworking, peaceful, good, clever, hostile, deceitful, lazy, violent, bad e stupid*), e as opções de resposta, “verdadeiro” e “falso”. O critério utilizado foi de, pelo menos, 80% de acertos em cada bloco, e uma latência média de resposta de até dois segundos em cada tentativa.

Após fazer o IRAP, os participantes respondiam a quatro medidas explícitas: as escalas de discriminação (DS) e diversidade (DV), escalas de diferencial semântico (SDS), e “*feeling thermometers*” (escalas avaliativas nas quais o participante indica o quão positiva ou negativa é sua atitude com relação a determinado estímulo a partir de graus, sendo consideradas atitudes negativas as com um grau baixo – e.g., 30 em uma escala de 0 a 100 – e atitudes positivas as com um grau alto – e.g., 90 na mesma escala supracitada). No IRAP, os participantes negros demonstraram vieses pró-branco e pró-negro em todos os quatro tipos de tentativas, enquanto os participantes brancos demonstraram vieses pró-branco nos dois tipos de tentativas que envolviam o rótulo “pessoas brancas” e viés pró-negro nas tentativas do tipo “negro-positivo”. No tipo de tentativa “negro-negativo”, os participantes brancos mostraram um forte viés anti-negro. Não é fácil explicar porque os participantes brancos demonstraram ambos os vieses pró-negro e anti-negro, dependendo da relação testada. Uma explicação dada pelos experimentadores foi de que as respostas obtidas no instrumento podem ser afetadas por “vieses verbais genéricos inerentes na linguagem” (Power et al., 2017a, p. 9). Além disso, foi levantada a possibilidade de que suas respostas foram afetadas pelo estímulo antecedente (controle de estímulos), de modo que, nas tentativas em que as pessoas negras foram relacionadas com coisas negativas, as representações negativas, frequentes na mídia, de personagens negros, influenciaram o responder, e nas tentativas que relacionavam os negros com coisas positivas, exemplares positivos de indivíduos negros, também presentes na mídia, como Barack Obama e Nelson Mandela, influenciaram as respostas.

As escalas DS e DV mostraram vieses pró-negro tanto para os participantes negros quanto brancos. Nas escalas de diferencial semântico, todos os participantes demonstraram vieses positivos, embora os participantes negros tenham avaliado as pessoas negras e brancas como mais positivas que a avaliação feita pelos participantes brancos. Os *feeling thermometers* indicaram que os participantes negros avaliaram negros como mais positivos que brancos. Do mesmo modo, participantes brancos avaliaram brancos como mais positivos que negros.

Os pesquisadores encontraram poucas correlações entre os resultados obtidos no IRAP e nas medidas explícitas (seis das 30 investigadas). Os experimentadores também fizeram análises de regressão logística hierárquica para verificar se as medidas do IRAP teriam validade preditiva sobre qual grupo de participantes (negro ou branco) respondeu a medidas. Apesar das poucas correlações entre o IRAP e medidas explícitas, os autores concluíram que o IRAP apresentou uma validade preditiva muito superior se comparado às medidas explícitas.

Por fim, a pesquisa de Power et al. (2017b) teve como objetivo avaliar vieses raciais de irlandeses brancos no IRAP, com o diferencial de que medidas de eletroencefalograma (EEG) foram obtidas enquanto os participantes faziam o IRAP. Dezesseis participantes adultos, brancos e com idades entre 18 e 33 anos participaram da pesquisa. Os estímulos-alvo

utilizados foram as frases “Eu acho que pessoas NEGRAS são” e “Eu acho que pessoas BRANCAS são”, os rótulos foram seis atributos positivos (amigável, honesto, trabalhador, pacífico, bom e esperto - *friendly, honest, hardworking, peaceful, good e clever*) e seis negativos (hostil, enganador, preguiçoso, violento, mau e estúpido - *hostile, deceitful, lazy, violent, bad e stupid*), e as opções de resposta, “verdadeiro” e “falso”. Os critérios de aprendizagem requeridos foram uma latência média de até dois segundos e acurácia média de 80% ou mais em cada bloco de tentativas.

Os sinais do EEG foram obtidos a partir de 32 eletrodos posicionados de acordo com o sistema internacional 10-20, distribuídos por todas as áreas do cérebro (pré-frontal, frontal, temporal, parietal, occipital e central). Os dados do EEG de seis participantes foram excluídos devido a dificuldades de interpretação dos padrões (*excessive noise*). Os resultados mostraram que os participantes demonstraram vieses pró-branco e anti-negro, e as medidas de EEG das áreas frontais do cérebro mostraram padrões de atividade diferentes quando as tentativas eram consistentes, em comparação com as inconsistentes. Apesar de tais resultados serem difíceis de interpretar, devido à ausência de estudos que correlacionem os resultados do IRAP com tais medidas (apenas um estudo fez esse tipo de comparação – Barnes-Holmes, Hayden, Barnes-Holmes, & Stewart, 2008), os autores sugerem que o IRAP pode se tornar uma ferramenta potencialmente útil para a área de neurociência afetiva.

DISCUSSÃO

O objetivo dessa pesquisa foi revisar os estudos publicados em periódicos que utilizaram o IRAP para investigar vieses, estereótipos e preconceitos raciais na literatura analítico-comportamental. Visando Visando um melhor entendimento dos dados produzidos nestes estudos empregando o IRAP, separamos dois pontos importantes.

Coerência ou ausência de coerência dos resultados do IRAP com medidas explícitas

Com relação aos resultados obtidos no IRAP, de maneira geral, as pesquisas revisadas demonstraram a presença de vieses raciais pró-branco e/ou anti-negro para os participantes brancos (Barnes-Holmes, et al., 2010 – Estudo 2; Power, et al., 2017a; 2017b). Por outro lado, com relação às medidas explícitas, os participantes geralmente demonstraram vieses positivos para ambos os grupos, i.e., brancos e negros (Barnes-Holmes, et al., 2010; Power et al., 2017a), ou até vieses mais positivos para negros do que para brancos (Barnes-Holmes, et al., 2010).

A coerência entre as medidas implícitas e explícitas foi avaliada em quatro dos sete estudos. Em dois deles, houve coerência entre os resultados obtidos com as medidas implícitas e explícitas (Mizael et al., 2016 e Power et al., 2017a), apesar de que no segundo estudo (Power et al., 2017a) tal coerência foi parcial, estando presente nos resultados dos participantes negros e parcialmente nos resultados dos participantes brancos.

Na literatura de instrumentos de avaliação implícita, sempre é salientado que este tipo de procedimento pode produzir dados potencialmente mais precisos (e.g. Hughes et al., 2011). Assim, em uma avaliação crítica deste ponto devemos considerar que a coerência ou ausência de coerência entre as medidas implícitas e explícitas deverá levar em consideração o histórico comportamental dos participantes. Sendo assim, a coerência não é necessariamente desejável como também a ausência dela não é necessariamente ruim. Nesse ponto vale ressaltar os mo-

delos explicativos já presentes na literatura para interpretação dos dados do IRAP, o modelo REC (Barnes-Holmes et al., 2010) e o modelo DAARRE (*Differential Arbitrarily Applicable Relational Responding Effects*), ou, em português, Efeitos diferenciais do responder relacional arbitrariamente aplicado - Finn, Barnes-Holmes, & McEntaggart, 2018).

De acordo com o modelo REC, a coerência entre as respostas obtidas com essas medidas vai depender do responder relacional de um indivíduo e de suas respostas breves e imediatas. Nas palavras de Barnes-Holmes et al. (2010): “respostas avaliativas imediatas ou automáticas podem ou não concordar com o responder relacional subsequente. Quando há coerência, medidas implícitas e explícitas serão tipicamente coerentes, mas quando não há coerência, as medidas irão tipicamente divergir” (p. 60; tradução nossa). Assim, quando observamos as explicações baseadas no modelo REC, está sendo destacada especialmente a história relacional dos indivíduos com os estímulos apresentados. Neste caso específico do preconceito racial, a forma como estas relações serão estabelecidas depende tanto de práticas culturais e também de experiências pessoais – que podem afetar muitas vezes a força e rigidez destas relações.

Assim, a coerência observada no primeiro dos dois estudos em que ela ocorreu (Mizael et al., 2016) possivelmente pode ter revelado características de uma história relacional ainda curta (os participantes eram crianças), o que possibilitou, possivelmente, que as relações treinadas durante o procedimento afetassem o responder destas crianças, o que é coerente com a proposta do modelo REC. Assim, o caráter socialmente sensível das tarefas também é algo que pode ter relação com a presença ou não de consistência entre os dados das medidas implícitas e explícitas. Desse modo, o fato de a pesquisa de Mizael et al. (2016) ter tido como participantes crianças que provavelmente tiveram uma menor história de repreensão consequente à emissão de respostas verbais “racistas”, além de haver uma maior aceitação desse tipo de comentário por crianças, pode auxiliar a entender a coerência obtida entre as medidas.

O modelo DAARRE, por outro lado, se direcionará para outros fatores além da rede relacional possivelmente estabelecida entre o rótulo e o alvo. O modelo DAARRE irá também atentar para as características funcionais dos estímulos e verificar potenciais coerências entre as funções dos elementos presentes no IRAP, tanto a relação do alvo e do rótulo, como também a relação deles com as opções de resposta, que serão chamados neste tipo de análise de Indicadores de Coerência Relacional (RCI, em inglês, *Relational Coherence Indicators*). Esses RCI permitirão explicar as tentativas pelos dois tipos de dicas contextuais (funcionais e relacionais) levando em consideração a coerência de suas funções. Por esse modelo, por exemplo, podemos entender a inconsistência que aconteceu entre os dados explícitos e implícitos no estudo de Power et al. (2017a), em que os participantes brancos, na medida explícita, responderam com um viés pró negro. Como, no IRAP, a tentativa “negro-positivo” tinha provavelmente baixa coerência graças a relação entre negro + positivo e negro + verdadeiro, isso acabou evocando, em uma frequência maior, a resposta falso (para maiores informações sobre esse tipo de análise, ver Finn et al., 2018).

Confiabilidade dos resultados obtidos a partir do IRAP

Um dado importante obtido na pesquisa de Drake et al. (2015) foram os resultados distintos nas duas aplicações do IRAP, feitas com os mesmos participantes, na qual houve uma diminuição dos vieses a favor do próprio grupo e aumento dos vieses a favor do outro grupo. Uma vez que o IRAP é um instrumento que mede relações consistentes e inconsistentes com a história de vida de um indivíduo, mudanças desse tipo não deveriam ocorrer sem algum

tipo de manipulação experimental. Assim, uma vez que, neste estudo, as duas aplicações ocorreram sem qualquer tipo de manipulação, esses resultados podem levantar questões sobre teste-reteste ou confiabilidade dos resultados. Por outro lado, diversas medidas novas têm sido empregadas na utilização deste instrumento graças a observações sucessivas de dados de difícil interpretação como esse. Recentemente, na última atualização do software do IRAP (GO-IRAP), é possível notar níveis diferentes de critérios de acerto (*Block level x Trial-type level*). Isto porque os desenvolvedores do programa notaram que mesmo obtendo mais que 90% de acertos em um bloco (ou seja, na média dos quatro tipos de tentativas), isso pode indicar índices altos de acertos em três tipos de tentativas e um muito baixo na quarta (e.g., 100%, 100%, 100%, 60%, média 90%). Para sanar este problema, o programa apresenta como opção o critério nível de tipo de tentativa (*Trial-type level*), ou seja, o critério estabelecido tem que ser atingido em todos os quatro tipos de tentativa, e não mais somente no bloco, independentemente do desempenho em cada tipo de tentativa. Dessa forma, é impossível obter um valor mais baixo que o exigido em um tipo de tentativa específica e seguir para a fase de testes.

Como se trata de um instrumento em constante desenvolvimento (tanto de versões como de modelos explicativos) e o seu objeto de estudo é muito complexo – isto é, o responder relacional arbitrariamente aplicável, devemos aguardar ainda os impactos desses novos critérios para poder afirmar com mais segurança sobre a fidedignidade de dados nos critérios antigos. Além disso, alguns estudos que apresentaram resultados de difícil interpretação, que tem como participantes adultos e utilizaram critérios muito baixos de acurácia (e.g. Drake et al., 2010, em que o critério foi uma média de 65% de acertos em cada bloco) dentro das possibilidades novas de interpretação, poderiam indicar dificuldades potenciais em elementos da etapa de treino do IRAP (para maiores informações sobre esses critérios, ver o Manual do GO-IRAP)⁴. Por fim, para verificar esse aspecto da fidedignidade dos dados do IRAP, particularmente a possibilidade de aplicações repetidas com os mesmos participantes, empregando o instrumento como uma medida de pré e pós-teste, ou mesmo observando a estabilidade desta medida, maiores investigações são necessárias para podermos fazer afirmações mais precisas.

Recentemente, um novo critério de análise, o Índice Probabilístico do IRAP (PI-IRAP, de Schryver, Hussey, Neve, Cartwright, & Barnes-Holmes, 2018) tem sido utilizado. Este permite potencialmente a investigação de dados individuais do IRAP, ao se calcular a probabilidade de cada participante emitir respostas coerentes ou incoerentes nos diferentes tipos de tentativas. Como também se trata de uma inovação do instrumento, este tipo de índice pode permitir a reavaliação de dados antigos e dar mais informações sobre a precisão e estabilidade dos resultados do IRAP.

Os estudos revisados têm evidenciado uma preocupação importante da área, que se refere ao estudo de tópicos socialmente sensíveis de maneira mais fidedigna. As pesquisas encontradas mostraram a presença de vieses negativos para pessoas negras, um dado muitas vezes não encontrado nas pesquisas de autorrelato, ou seja, com o uso de medidas explícitas. Apesar desse tipo de resultado ser importante socialmente e cientificamente falando, uma vez que justifica políticas afirmativas e sanções por injúria racial e outros tipos de crimes e delitos, entre outras medidas, outras pesquisas utilizando o IRAP (e outros instrumentos, como o FAST) podem aumentar o escopo e alcance da análise do comportamento no estudo de questões sociais. Algumas sugestões são a inclusão de análises demográficas (e.g., identidade de gênero

4) Disponível em <https://go-ift.com/wp-content/uploads/2018/01/GO-IRAP-manual-v1.1.pdf>

e status socioeconômico dos participantes) com o intuito de verificar possíveis diferenças nos resultados, a depender de diferentes marcadores sociais, como os supracitados; o uso de medidas explícitas similares ou idênticas a de estudos prévios, permitindo a comparação com pesquisas anteriores; a criação de um protocolo com o intuito de reduzir vieses raciais negativos de participantes que fazem o IRAP e demonstram vieses raciais negativos para algum grupo; e a verificação de possíveis relações entre os resultados do IRAP e tarefas comportamentais que ilustrem comportamentos discriminatórios. Em um fim último, instrumentos de medidas implícitas como o IRAP possuem um potencial para revelar detalhes sobre o histórico de vida dos participantes, emitindo respostas que seriam mais difíceis de serem elaboradas e, dessa forma, sensíveis a sanções sociais.

Existem ainda muitos pontos a serem mais desenvolvidos em experimentos futuros: o modelo DAARRE, por exemplo foi recentemente apresentado, e ele impacta nas escolhas de opções de resposta (como RCI), mostrando que mesmo as escolhas destas podem afetar o resultado do instrumento; outra variável que necessitará de investigações são os critérios de precisão presentes na versão mais nova do programa, pois ainda não foram feitas pesquisas investigando seu potencial, apesar de teoricamente o nível de tipo de tentativas permitir fazer uma inferência de que esse tipo de análise trará uma maior precisão dos dados observados⁵. A própria possibilidade de usar o IRAP como um instrumento de análise em dois momentos, depende de muitos estudos verificando a estabilidade da avaliação que ele apresenta. Em suma, o IRAP pode ser considerado um instrumento que, apesar de estar em desenvolvimento, possui resultados promissores que podem ser úteis⁶ e, inclusive, mais fidedignos que algumas medidas explícitas, especialmente nos contextos de pesquisa com temas socialmente sensíveis, como investigações sobre estereótipos e preconceitos.

REFERÊNCIAS

Barnes-Holmes, D., Barnes-Holmes, Y., Power, P., Hayden, E., Milne, R., & Stewart, I. (2006). Do you know what you really believe? Developing the Implicit Relational Assessment Procedure (IRAP) as a direct measure of implicit beliefs. *The Irish Psychologist*, 32(7), 169-177. Recuperado de: https://contextualscience.org/publications/do_you_really_know_what_you_believe_developing_the_irap_as_a_direct_measure

5) Os estudiosos da Teoria das Molduras Relacionais recentemente têm proposto outros modelos para permitir a interpretação de respostas relacionais derivadas. O modelo chamado Multi-Dimensional Multi-Nível (MDML, Barnes-Holmes, Barnes-Holmes, Luciano, & McEnteggart, 2017), permite uma interpretação do RRAA de forma organizada, distribuindo características relacionais desse responder de forma dinâmica entre cinco níveis e quatro dimensões (Coerência, Flexibilidade, Derivação e Complexidade) que permitem uma análise bastante precisa do repertório comportamental complexo de seres humanos. Esse modelo, apesar de recente, tem se mostrado de grande relevância para o entendimento das propriedades do responder relacional derivado, e inclusive, recebeu uma atualização. Essa atualização considera tanto elementos relacionais (relacionar) como também funcionais (evocar e orientar), passando a ser chamado de modelo Hiper-Dimensional Multi-Nível (Bond, Dowling, Barnes-Holmes, & Barnes-Holmes, 2018).

6) Para algumas críticas ao instrumento, ver, por exemplo, Cummins, J., Roche, B., Tyndall, I., & Cartwright, A. (2018). The relationship between differential stimulus relatedness and implicit test effect sizes. *Journal of the Experimental Analysis of Behavior*, 118, 24-38. doi: 10.1002/jeab.437 e Cartwright, A., Roche, B., Gogarty, M., O'Reilly, A. & Stewart, I. (2016). Using a modified Function Acquisition Speed Test (FAST) for assessing implicit gender stereotypes. *The Psychological Record*, 66, 223-233. doi: 10.1007/s40732-016-0164-5

- Barnes-Holmes, D., Hayden, E., Barnes-Holmes, Y., & Stewart, I. (2008). The Implicit Relational Assessment Procedure (IRAP) as a response-time and event-related-potentials methodology for testing natural verbal relations: A preliminary study. *The Psychological Record*, *58*, 497–516. doi: 10.1007/BF03395634
- Barnes-Holmes, D., Murphy, A., 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*, 57-66. doi: 10.1007/BF03395694
- 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*, 527-542. doi: 10.1007/BF03395726
- Barnes-Holmes, Y., Luciano, C., & McEnteggart, C. (2017). From the IRAP and REC model to a multi-dimensional multi-level framework for analyzing the dynamics of arbitrarily applicable relational responding. *Journal of Contextual Behavioral Science*, *6*(4), 434-445. doi: 10.1016/j.jcbs.2017.08.001
- Bond, F., Dowling, J., Barnes-Holmes, D., & Barnes-Holmes, Y. (2018) Alive and Thriving: a functional contextual, RFT-consistent model for achieving meaning, vitality, and connection amongst people. Apresentação oral no *1 Chester Contextual Behavioural Science Research Colloquium*.
- Boysen, G. A., Vogel, D. L., & Madon, S. (2006). A public versus private administration of the Implicit Association Test. *European Journal of Social Psychology*, *36*, 845-856. doi: 10.1002/ejsp.318
- Campbell, C., Barnes-Holmes, Y., Barnes-Holmes, D., & Stewart, I. (2011). Exploring screen presentations in the Implicit Relational Assessment Procedure (IRAP). *International Journal of Psychology and Psychological Therapy*, *11*(3), 377-388. Recuperado de: <https://www.ijpsy.com/volumen11/num3/305/exploring-screen-presentations-in-the-implicit-EN.pdf>
- Cartwright, A., Hussey, I., Roche, B., Dunne, J., & Murphy, C. (2016). An investigation into relationship between the gender binary and occupational discrimination using the Implicit Relational Assessment Procedure. *The Psychological Record*, *67*(1), 121-130. doi: 10.1007/s40732-016-0212-1
- De Houwer, J. (2003). The extrinsic affective Simon Task. *Experimental Psychology*, *50*, 77-85. doi: 10.1027//1618-3169.50.2.77
- De Houwer, J. (2006). What are implicit attitudes and why are we using them? In: R. W. Wiers, & A. W. Stacy (Eds.). *The Handbook of Implicit Cognition and Addiction* (pp. 11-28). Thousand Oaks, CA: Sage.
- de Houwer, J., Heider, N., Spruyt, A., Roets, A., & Hughes, S. (2015). The Relational Responding Task: Toward a new implicit measure of beliefs. *Frontiers in Psychology*, *6*, 1-9. doi: 10.3389/fpsyg.2015.00319
- de Schryver, M., Hussey, I., Neve, J., Cartwright, A., & Barnes-Holmes, D. (2018). The PI IRAP: An alternative scoring algorithm for the IRAP using a probabilistic semiparametric effect size measure. *Journal of Contextual Behavioral Science*, *7*, 97-103. doi: 10.1016/j.jcbs.2018.01.001

- Drake, C. E., Kellum, K. K., Wilson, K. G., Luoma, J. B., Weinstein, J. H., & Adams, C. H. (2010). Examining the Implicit Relational Assessment Procedure: Four preliminary studies. *The Psychological Record, 60*, 81-100. doi: 10.1007/BF03395695
- Drake, C. E., Kramer, S., Sain, T., Swiatek, R., Kohn, K., & Murphy, M. (2015). Exploring the reliability and convergent validity of implicit racial evaluations. *Behavior and Social Issues, 24*, 68-87. doi:10.5210/bsi.v.24i0.5496.
- Farrel, L., & McHugh, L. (2017). Examining gender-STEM bias among STEM and non-STEM students using the Implicit Relational Assessment Procedure (IRAP). *Journal of Contextual Behavioral Science, 6*(1), 80-90. doi: 10.1016/j.jcbs.2017.02.001
- Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: A bona fide pipeline? *Journal of Personality and Social Psychology, 69*, 1013-1027. doi: 10.1037/0022-3514.69.6.1013
- Finn, M., Barnes-Holmes, D., & McEnteggart, C. (2018). Exploring the single-trial-type-dominance-effect in the IRAP: Developing a Differential Arbitrarily Applicable Relational Responding Effects (DAARRE) model. *The Psychological Record, 68*(1), 11-25. doi: 10.1007/s40732-017-0262-z
- Freedman, J. L., Carlsmith, J. M., & Sears, D. O. (1970). *Psicologia Social*. São Paulo: Cultrix. doi: 10.1037/0022-3514.69.6.1013
- Goethals, G. R., & Darley, J. (1977). Social comparison theory: An attributional approach. In J. M. Suls, & R. L. Miller (Eds.). *Social comparison processes: Theoretical and empirical perspectives* (pp. 259-278). Washington, DC: Hemisphere.
- Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Psychological Review, 102*, 4-27. doi: 10.1037/0033-295X.102.1.4
- 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*, 1464-1480. doi: 10.1037/0022-3514.74.6.1464
- Greenwald, A. G., Nosek, B. A., & Banaji, M. R. (2003). Understanding and using the Implicit Association Test: I. An improved scoring algorithm. *Journal of Personality and Social Psychology, 85*, 197-216. doi: 10.1037/0022-3514.85.2.197
- Hayes, S. C., Barnes-Holmes, D., & Roche, B. (2001). *Relational frame theory: A post-Skinnerian account of human language and cognition*. New York, NY: Kluwer Academic.
- Hughes, S., Barnes-Holmes, D., & de Houwer, J. (2011). The dominance of associative theorizing in implicit attitude research: Propositional and behavioral alternatives. *The Psychological Record, 61*, 45-496. doi: 10.1007/BF03395772
- Kohlsdorf, M., & Costa Junior, L. (2009). O autorrelato na pesquisa em psicologia da saúde: Desafios metodológicos. *Psicologia Argumento, 27*, 131-139.
- 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*, 451-462. doi: 10.1007/s40732-016-0185-0
- Nogueira, O. (2006). 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: Revista de Sociologia da USP, 19*(1), 287-308. doi: 10.1590/S0103-20702007000100015
- Nosek, B., & Banaji, M. R. (2001). The go/no-go association task. *Social Cognition, 19*, 625-666. doi: 10.1521/soco.19.6.625.20886

- 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. doi: 10.1007/BF03395817
- Paluck, E. L., & Green, D. (2009). Prejudice reduction: What works? A review and assessment of research and practice. *Annual Reviews of Psychology*, 60, 339-367. doi: 10.1146/annualrev.psych.60.110707.163607
- Power, P. M., Barnes-Holmes, D., Barnes-Holmes, Y., & Stewart, I. (2009). The Implicit Relational Assessment Procedure (IRAP) as a measure of implicit relative preferences: A first study. *The Psychological Record*, 59, 621-640. doi: 10.1007/BF03395684
- Power, P. M., Harte, C., Barnes-Holmes, D., & Barnes-Holmes, Y. (2017a). Exploring racial bias in a European country with a recent history of immigration of Black Africans. *The Psychological Record*, 67(3), 365-375. doi: 10.1007/s40732-017-0223-6
- Power, P. M., Harte, C., Barnes-Holmes, D., & Barnes-Holmes, Y. (2017b). Combining the Implicit Relational Assessment Procedure and the recording of event related potentials in the analysis of racial bias: A preliminary study. *The Psychological Record*, 67(4), 499-506. doi: 10.1007/s40732-017-0252-1
- Rabelo, L. Z., Bortoloti, R., & Souza, D. H. (2014). Dolls are for girls and not for boys: Evaluating the appropriateness of the IRAP for school-age children. *The Psychological Record*, 64(1), 71-77. doi: 10.1007/s40732-014-0006-2
- Richeson, J. A., & Sommers, S. R. (2015). Toward a social psychology of race and race relations for the twenty-first century. *Annual Review of Psychology*, 67, 18.1-18.25. doi: 10.1146/annurev-psych-010213-115115
- Ritzert, T. R., Anderson, L. M., Reilly, E. E., Gorrell, S., Forsyth, J. P., & Anderson, D. A. (2016). Assessment of weight/shape implicit bias related to attractiveness, fear, and disgust. *The Psychological Record*, 66(3), 405-417. doi: 10.1007/s40732-016-0181-4
- Skinner, B. F. (1953). *Science and Human Behavior*. New York: McMillan. Disponível em www.bfskinner.org/newtestsite/wp-content/uploads/.../ScienceHumanBehavior.pdf
- Schwarz, N., & Bohner, G. (2001). The construction of attitudes. In A. Tesser, & N. Schwarz (Eds.), *Blackwell Handbook of Social Psychology: Intrapersonal processes*. (pp. 436-457). Oxford, UK: Blackwell.
- Tajfel, H. (1982). Social psychology of intergroup relations. *Annual Review of Psychology*, 33, 1-39. doi: 10.1146/annurev.ps.33.020182.000245
- Vahey, N. A., Nicholson, E., & Barnes-Holmes, D. (2015). A meta-analysis of criterion effects for the Implicit Relational Assessment Procedure (IRAP) in the clinical domain. *Journal of Behaviour Therapy and Experimental Psychiatry*, 48, 59-65. doi: 10.1016/j.jbtep.2015.01.004

Received: September 15, 2018

Accepted: April 30, 2019