

Laboratório de Psicologia, 10(2): 235-249 (2012)
© 2012, I.S.P.A.

Paradigma: Teste de associação implícita

Manuel Oliveira
Mariana P. Miranda
ISPA – Instituto Universitário

Resumo

O paradigma do teste de associação implícita (Implicit Association Test ou IAT), proposto por Greenwald, McGhee e Schwartz (1998) como instrumento para estudar a força das relações associativas entre conceitos e representações cognitivas, tem sido utilizado como medida implícita de eleição para aceder aos mais variados constructos tais como atitudes, crenças ou estereótipos, em diversos contextos de investigação (e.g., clínica, organizacional, política) para além da cognição social. O presente artigo descreve em pormenor os aspectos metodológicos associados ao IAT, informando ainda sobre as variáveis que moderam os efeitos IAT e que devem ser consideradas no planeamento dos estudos. São apresentadas algumas das explicações teóricas para os mecanismos subjacentes ao efeito IAT que têm sido propostas na literatura, e ainda resumidas as principais críticas de que o IAT tem sido alvo.

Palavras-chave: IAT, Medidas implícitas, Paradigma, Teste de Associação Implícita.

Abstract

The implicit association test paradigm (IAT), introduced by Greenwald, McGhee and Schwartz (1998) as an instrument to study the strength of associative links between concepts and cognitive representations, became, in recent years, a measure of choice to access constructs such as attitudes, beliefs or stereotypes, not only within the fields of social psychology and social cognition, but in several other research domains (e.g., clinical, organizational, politics) as well. The present article describes in detail the methodological aspects intrinsic to the IAT, and provides information about variables known to moderate the IAT's effects, which need to be carefully considered while planning studies. We also review some of the most relevant theoretical explanations to the mechanisms underlying the IAT effect and present the main criticism directed at the paradigm.

Key-words: IAT, Implicit Association Test, Implicit measures, Paradigm.

A correspondência relativa a este artigo deverá ser enviada para: Manuel Oliveira, ISPA – Instituto Universitário, Rua Jardim do Tabaco, 34, 1149-041 Lisboa; E-mail: manueljbo@gmail.com

A par do desenvolvimento da disciplina de cognição social surgiu, durante a última década do século XX, um crescente interesse pelos processos mentais inconscientes que estimulou o desenvolvimento de novas medidas de cognição implícita. De entre estas medidas propomos individualizar, neste artigo, o Teste de Associação Implícita (designado na literatura por *Implicit Association Test* ou IAT), apresentado no artigo seminal¹ de Greenwald, McGhee e Schwartz (1998), intitulado “Measuring individual differences in implicit cognition: The implicit association test”. Desde a sua publicação, o número de aplicações do IAT como medida implícita cresceu significativamente na literatura, estabelecendo-o como um novo paradigma não só em cognição social implícita, como também noutros contextos de investigação (e.g., clínico, desenvolvimento, marketing – para uma revisão ver Lane, Banaji, Nosek, & Greenwald, 2007).

Alinhado com as teorias associativas da aprendizagem e das representações, o IAT assenta na assunção de que é mais fácil dar uma resposta comportamental (e.g., pressionar uma tecla) quando esta reflecte dois conceitos associados entre si associados entre si, do que quando os mesmos exigem respostas diferentes (Greenwald et al., 1998). O efeito associado ao IAT observa-se assim quando num contexto comparativo entre dois pares de conceitos os tempos de reacção (TR) são significativamente mais rápidos para uma combinação dos conceitos do que para a sua combinação inversa.

O estabelecimento do efeito: Os estudos originais

Com o objectivo de introduzir e demonstrar a potencial utilidade de um novo método – o IAT – no diagnóstico de estruturas associativas socialmente significantes, Greenwald e colaboradores (1998) conduziram três estudos onde testaram a proficiência do IAT na mensuração de associações avaliativas subjacentes a atitudes implícitas².

O primeiro estudo testou a assunção principal do IAT de que as associações podem ser reveladas através do mapeamento num único par de respostas comportamentais de duas tarefas de discriminação. Para o efeito, construíram dois IATs com, por um lado, categorias correspondentes a objectos atitudinais familiares e de fraca sensibilidade social (viz. Flores-Insectos, Armas-Instrumentos Musicais) e, por outro, categorias que representavam conceitos avaliativos (Agradável-Desagradável). Os resultados demonstraram menores TR para combinações congruentes entre categorias (e.g., Flores e Agradável *vs.* Insectos e Desagradável) do que para combinações incongruentes (e.g., Flores e Desagradável *vs.* Insectos e Agradável), sustentando a exequibilidade da assunção central do IAT. Destes resultados foi inferida uma maior preferência relativamente a flores e instrumentos musicais, comparativamente a insectos e armas, respectivamente. Foi ainda neste estudo que se obtiveram as primeiras de uma série de baixas correlações entre medidas implícitas e explícitas (média $r=.19$), tendo-se adicionalmente observado uma dimensão dos efeitos superior nas medidas explícitas (média $d=1.68$) do que nos IATs (média $d=1.50$).

No segundo e terceiro estudos, os autores procuraram aceder a atitudes implícitas de maior sensibilidade social, recorrendo a categorias-alvo relativas à pertença grupal dos participantes. Assim, no segundo estudo utilizaram-se as categorias-alvo Japonês *versus* Coreano num IAT aplicado a participantes membros destes dois grupos, e no terceiro estudo, testou-se um IAT com as categorias-alvo Branco *versus* Negro numa amostra de membros do primeiro grupo. Os efeitos IAT dos estudos demonstraram um padrão de preferência pelo endogrupo, uma vez que os participantes responderam mais rapidamente aos blocos onde o seu endogrupo era combinado com a categoria com atributos positivos. Comparativamente ao primeiro estudo, nestes últimos a dimensão dos efeitos foi

¹ O artigo conta actualmente com cerca de 3810 citações segundo uma pesquisa no *Google Académico*.

² Greenwald e Banaji (1995) definem atitude implícita como “traços de experiências passadas não identificáveis por meio de introspecção, que influenciam julgamentos avaliativos de determinado objecto”.

substancialmente maior nos IATs do que nas medidas explícitas, tendo este padrão sido mais acentuado no terceiro estudo.

As evidências fornecidas pelos três estudos (a) confirmaram a utilidade do IAT no domínio das associações avaliativas entre pares semânticos ou categorias sociais e (2) sugerem que o IAT pode ser resistente a estratégias de auto-apresentação no acesso a associações pessoal ou socialmente sensíveis.

A definição do paradigma

Apesar de relativamente recente, o IAT tem sido aplicado e analisado em numerosos estudos e diversos contextos de investigação. Os aspectos transversais a todos eles desdobram-se num conjunto de procedimentos e delineamentos que aqui são definidos como o paradigma experimental do teste de associação implícita.

Convém ainda salientar que, após a sua primeira versão, proposta e testada no artigo seminal, o IAT foi submetido a um conjunto de testes que procuraram refinar as suas qualidades psicométricas. Estes esforços resultaram num IAT aperfeiçoado (para revisão ver Greenwald, Nosek, & Banaji, 2003), sendo nesta última versão que o presente artigo se foca.

Participantes

O IAT é reconhecido na literatura pela sua capacidade de produzir efeitos de grande dimensão a partir de amostras relativamente reduzidas (e.g., 20 a 30 participantes, ver Greenwald et al., 1998). O tamanho da amostra recomendável pode no entanto variar consoante o procedimento de cálculo utilizado ou análises adicionais (e.g., correlações IAT-medidas explícitas, ver p. 214, Greenwald et al., 2003).

Variáveis Independentes

Categorias do IAT. O IAT permite a utilização de quaisquer dois pares de categorias, cujas forças associativas se pretendam comparar (A1 vs. A2, categorias-alvo; e B1 vs. B2, categorias-atributo). Os conceitos (rótulos) utilizados em cada par devem ser antagónicos e facilmente identificáveis (e.g., Negro vs. Branco e Bom vs. Mau; Flores vs. Insectos e Agradável vs. Desagradável) (Nosek, Greenwald, & Banaji, 2007). É importante ter em conta que as palavras seleccionadas como conceitos determinam o enquadramento interpretativo dos estímulos que representam no IAT, assim como a interpretação do efeito observado (Mitchell, Nosek, & Banaji, 2003). Esta variável é necessariamente manipulada intra-sujeitos.

Material

Estímulos associados a cada categoria. A cada categoria corresponde uma lista de estímulos pré-testados que podem assumir a modalidade de palavra, imagem, som ou símbolo, consoante os objectivos do estudo (Nosek et al., 2007). É também possível utilizar diferentes modalidades de estímulo para cada par de categorias (e.g., imagens para Homem-Mulher, palavras para Bom-Mau). É essencial que os estímulos que se associam a cada categoria (a) sejam facilmente

identificados como membros exclusivos da sua categoria superordenada (e.g., “amor” para Bom; “guerra” para Mau) e (b) que o seu conjunto represente integralmente, ou tanto quanto possível, todos os aspectos do conceito ou atitude (e.g., não utilizar apenas lésbicas para a categoria *gay*) (Mitchell et al., 2003). Note-se ainda que o grau em que os estímulos são representativos das categorias se encontra positivamente correlacionado com a magnitude do efeito IAT (Feroni & Bel-Bahar, 2010). A partir de dois itens, a quantidade de itens por categoria não parece afectar a magnitude e consistência interna dos efeitos IAT (ver p. 174, Nosek, Greenwald, & Banaji, 2005).

Medidas Dependentes

Tempos de reacção. O procedimento do IAT baseia-se na medição computadorizada do intervalo de tempo que os participantes demoram a classificar cada estímulo na categoria-alvo que lhe corresponde. Por exigir rapidez de resposta, o tempo é tipicamente medido à escala dos milissegundos (ms) pelo programa informático (e.g., Inquisit, E-Prime) seleccionado para correr o IAT.

Frequência de erros. Quando a associação entre as categorias codificadas para a mesma resposta se assume ser inexistente ou fraca, espera-se um aumento na probabilidade de cometer erros na classificação dos estímulos em consequência da competição de resposta. Trata-se assim de uma medida complementar aos TR.

Procedimento

O procedimento do IAT engloba um total de sete passos (blocos), descritos de seguida. Os passos 4 e 7 correspondem aos blocos críticos, de onde se extraem os dados de interesse do IAT.

Bloco 1 – Aprendizagem do par de conceitos-alvo. No primeiro bloco são apresentados dois conceitos-alvo (e.g., Flores vs. Insectos), um localizado no canto superior esquerdo do ecrã e o outro no direito. A cada conceito-alvo corresponde uma de duas teclas, que os sujeitos vão pressionando à medida que surgem, de forma aleatória, estímulos associados a cada um dos conceitos no centro do ecrã (e.g., centopeia, tília, vespa, rosa). Tipicamente são utilizadas as teclas “E” e “I” do teclado, podendo também utilizar-se os botões de uma caixa de resposta. A tarefa dos sujeitos consiste em categorizar o mais rápido e correctamente possível cada estímulo.

Bloco 2 – Aprendizagem do par de conceitos-atributo. Segue-se um bloco idêntico ao primeiro com a excepção de os conceitos iniciais serem substituídos por dois novos conceitos (e.g., Agradável vs. Desagradável), geralmente designados por conceitos-atributo, e respectivos estímulos (e.g., paz, guerra, alegria, dor).

Bloco 3 – Treino da primeira tarefa combinada. Os conceitos-alvo e conceitos-atributo, apresentados nos dois primeiros blocos, aparecem agora simultaneamente no ecrã de forma combinada, de modo congruente (e.g., Flores e Agradável vs. Insectos e Desagradável) ou incongruente (e.g., Flores e Desagradável vs. Insectos e Agradável)³. A ordem de apresentação da tarefa combinada – congruente-primeiro ou incongruente-primeiro – deve ser contrabalançada. Este bloco serve para treinar os participantes na tarefa que é, por si, mais complexa do que nos blocos anteriores.

Bloco 4 – Teste crítico da primeira tarefa combinada. Após uma breve pausa, os participantes realizam, neste bloco, exactamente a mesma tarefa do bloco anterior.

³ A assunção de congruência ou incongruência dos conceitos que partilham a mesma resposta comportamental, assenta nas hipóteses de cada estudo em particular. Conceitos congruentes assumem-me como mais fortemente associados do que conceitos incongruentes.

Bloco 5 – Aprendizagem da inversão espacial dos conceitos. O quinto bloco apresenta os mesmos conceitos-alvo e estímulos do bloco inicial, invertendo a posição em que apareciam anteriormente (e.g., Insectos vs. Flores).

Bloco 6 – Treino da segunda tarefa combinada. Este bloco é idêntico ao terceiro com a exceção de exibir uma combinação inversa de compatibilidade entre os dois pares de conceitos, resultante da troca de posição dos conceitos-alvo, incongruente aqui no caso de congruente nas tarefas 3 e 4 ou o reverso.

Bloco 7 – Teste crítico da segunda tarefa combinada. Por fim, e após uma breve pausa, os participantes realizam exactamente a mesma tarefa do bloco anterior.

Análises estatísticas

Cálculo da medida D. Também os procedimentos de análise do IAT evoluíram desde a sua proposta inicial em 1998. Neste artigo apresentamos aquele que foi considerado, após extensa recolha empírica, o melhor candidato a algoritmo de análise dos dados do IAT com base em critérios psicométricos (e.g., sensibilidade a influências conhecidas, consistência interna, correlações com medidas explícitas, resistência à influência de variáveis externas) (Greenwald et al., 2003). O algoritmo designa-se por algoritmo D. A utilização deste algoritmo minimiza (a) a correlação entre os efeitos IAT e as médias individuais de TR, (b) o efeito da ordem dos blocos do IAT, e (c) os efeitos de uma exposição prévia ao teste, ao mesmo tempo que maximiza a correlação entre o IAT e as medidas explícitas e mantém uma boa consistência interna (Lane et al., 2007).

O algoritmo recomendado por Greenwald e colaboradores (2003), para delineamentos de IATs onde os participantes têm de corrigir as respostas erradas antes de prosseguir para o ensaio seguinte tem os seguintes passos:

- 1 – Usar os dados dos blocos 3, 4, 6 e 7 (blocos críticos);
- 2 – Eliminar ensaios com TR superiores a 10000 ms;
- 3 – Eliminar participantes para quem 10% dos ensaios apresentem TR inferiores a 300ms;
- 4 – Calcular um desvio-padrão de todos os ensaios dos blocos 3 e 6. Fazer o mesmo para os blocos 4 e 7;
- 5 – Calcular a média dos TR de todos os ensaios de cada um dos quatro blocos (3,4,6,7);
- 6 – Calcular a diferença entre os blocos 3 e 6 (blocos de treino) e entre os blocos 6 e 7 (blocos críticos), subtraindo-se o bloco congruente do bloco incongruente (Bloco incongruente – Bloco congruente);
- 7 – Dividir cada diferença pelo seu desvio-padrão associado (Passo 4);
- 8 – Calcular a média dos dois quocientes obtidos no Passo 7.

No caso de o IAT não exigir que os participantes corrijam as respostas antes de prosseguirem para o ensaio seguinte, recomenda-se (a) que apenas se utilizem os ensaios das respostas correctas nos Passos 3 e 4, e (b) a substituição de cada TR das respostas erradas pelo valor da média de todos os ensaios correctos do bloco somado a uma penalização de 600 ms, antes de se completar o Passo 5. Este último algoritmo é conhecido na literatura como o algoritmo D600 (e.g., De Hower & De Bruycker, 2007). Uma das características estruturais do IAT é a sua natureza relativa, variando os efeitos IAT em função das combinações entre os dois pares de conceitos. Como tal, deve-se ter em conta que as respostas aos estímulos de um IAT se inserem numa estrutura de comparação que se

reflece em cada um dos ensaios. Deste modo, os efeitos IAT devem ser calculados espelhando essa comparação, e não basear-se na análise independente dos TR de apenas uma categoria (ver p. 92, Lane et al., 2007; Nosek et al., 2005).

A pontuação de um IAT corresponde à dimensão do efeito obtido, cujo cálculo segue a mesma fórmula do *d de Cohen*, com a excepção do desvio-padrão incluir todos os ensaios de ambos os blocos críticos do IAT (ver cálculo da medida D). Para não se confundir com o *d de Cohen*, a pontuação de um IAT é designada por medida D ou *D-score* (ver p. 201, Greenwald et al., 2003). Os valores de D oscilam entre -2 e +2, servindo como indicadores da força relativa de associação entre os conceitos. Assume-se que um D positivo significativamente diferente de zero, é indicador de uma maior associação entre os conceitos do bloco onde estes se consideram combinados de forma congruente, do que no bloco onde estes se combinam de forma incongruente. Por outro lado, um D negativo indicará uma maior associação entre os conceitos da combinação considerada incongruente.

Análises complementares. Complementarmente à medida D, é ainda comum (a) analisar transformações logarítmicas das médias de TR ($\log TR$)⁴ dos blocos críticos com recurso a testes-*t* ou ANOVAs de medições repetidas (consoante o número de variáveis consideradas), (b) analisar a validade convergente entre o IAT e uma ou várias medidas explícitas, através do cálculo da correlação entre a medida D do IAT e um índice equivalente ao D computado com os valores das medidas explícitas, e (c) calcular a consistência interna (e.g., método da divisão em metades). É também comum, embora facultativo, complementar a análise dos TR com uma análise da frequência de erros em cada bloco. Esta análise permite confirmar a correspondência inequívoca entre cada lista de estímulos e a sua respectiva categoria (e.g., dois primeiros blocos), ou ainda complementar os resultados dos TR nos blocos críticos, ao informar sobre a taxa de erros obtida em cada um.

Variações do paradigma

Os esforços movidos no sentido do refinamento de algumas das limitações e/ou ambiguidades associadas ao IAT, no que concerne à interpretação dos efeitos obtidos ou a especificidades da amostra e contexto de aplicação, inspiraram a introdução de algumas variações que resultaram em versões alternativas à medida original. Estas variações, descritas de seguida, diferem do IAT original quer no formato de aplicação, quer na estrutura e procedimento.

IAT Papel-e-Lápis

Quanto ao formato de aplicação, surgiu uma versão de papel-e-lápis do IAT cuja principal vantagem consiste na aplicação do IAT em contextos onde a sua versão computarizada seja impraticável, permitindo a utilização de amostras não-tradicionais (e.g., banhistas na praia, transporte público). Apesar desta vantagem, a imprecisão de cronometragem, a que se junta o facto de ser necessário a utilização de itens omissos como medida dependente, conduzem à redução da magnitude dos efeitos, impedindo a versão *low-tech* de servir como substituto à computarizada (Lemm, Lane, Sattler, Khan, & Nosek, 2008).

Brief IAT (BIAT)

Esta versão do IAT foi desenvolvida com o objectivo de reduzir variações espontâneas relacionadas com possíveis estratégias procedimentais dos participantes (e.g., responder com base na saliência cognitiva de

⁴ Greenwald e colaboradores (1998) optaram por efectuar transformações logarítmicas aos TR de forma a estabilizar a variância da estatística a aplicar sobre os dados. A distribuição dos TR é tipicamente exponencial, o que inviabiliza o recurso a análises que possuam como requisito a homocedasticidade (e.g., ANOVAs). Os TR transformados logaritmicamente apresentam maior simetria nas caudas da distribuição, aproximando a sua distribuição à normal.

um par de categorias – ver modelo da assimetria figura-fundo na secção “Explicações Alternativas do Efeito IAT”). Comparativamente ao IAT, o BIAT inclui apenas os dois blocos de tarefas combinadas com instruções para os participantes se focarem em apenas duas das quatro categorias. Durante a tarefa de categorização, os participantes mantêm a sua atenção focada nesse par de categorias focais, pressionando uma tecla para “estímulos que correspondam a qualquer umas das categorias do par focal”, e outra tecla para “qualquer estímulo que não corresponda ao par focal”. A convergência entre os efeitos do BIAT e do IAT parece depender das categorias focais (Sriram & Greenwald, 2009).

Single Category IAT (SC-IAT)

Nesta variante do IAT utiliza-se apenas uma categoria-alvo e duas categorias-atributo. Aplicando dois SC-IAT é, deste modo possível ir além da interpretação necessariamente comparativa do efeito IAT, identificando qual a relação associativa entre cada categoria-alvo e as categorias-atributo. Torna-se assim possível identificar a qual das associações compatíveis (e.g., flores e agradável ou armas e desagradável) se deve o efeito IAT geral (Karpinski & Steinman, 2006).

Multifactor Trait IAT (MFT-IAT)

O MFT-IAT foi arquitectado por Greenwald (2005) com o propósito de aceder implicitamente a traços em estudos onde se utilizam inventários de personalidade como medidas explícitas. Os inventários de personalidade mais popularmente utilizados como o “*Big Five*” ou o “*Myers-Briggs Type Indicator*” incluem traços de personalidade centrais que incluem dois pólos de valência (e.g., Amabilidade: Amável vs. Não-amigável). Em estudos onde se pretenda aceder a associações implícitas entre o *self* (categorias-alvo: Eu vs. Outros) e os traços, convém garantir que as respostas num IAT não se baseiem na valência das categorias-atributo (e.g., Amável vs. Não-amigável), o que resultaria num IAT de auto-estima que não reflectiria as pretendidas associações entre o traço e o *self*, mas sim a avaliações do self. Por esta razão, contrariamente ao IAT original, apenas uma das categorias-atributo do MFT-IAT é relevante para o alvo, não sendo a segunda o seu oposto (e.g., categorias-alvo: Eu vs. Outros; categorias-atributo: Extraversão vs. Outros traços para além de extraversão ou intraversão). Desta forma, a valência das categorias-atributo e respectivos estímulos é homogénea (i.e., todos positivos, negativos ou neutros), evitando-se a possível estratégia procedimental.

Go/No-go Association Task (GNAT)

Esta técnica visa examinar processos automáticos face a uma única categoria-alvo sem o envolvimento de objectos complementares ou contrastantes. O GNAT tem replicado efeitos IAT associados a grupos de género e racializados. Uma das suas vantagens sobre o IAT reside na possibilidade de estudar categorias sem um oposto semântico óbvio (Nosek & Banaji, 2001).

Child-IAT

Trata-se de uma versão do IAT adaptada para amostras de crianças desenvolvida por Baron e Banaji (2006). Nesta versão as palavras-estímulo são substituídas por gravações de voz em consequência da variabilidade de compreensão de texto entre as crianças, e, sempre que os estímulos visuais sejam correspondentes a pessoas, são utilizadas faces de crianças. O Child-IAT é apresentado às crianças como um videojogo, e elas respondem através de dois controlos, cada um com um botão relativamente grande (e.g., 7,6 cm, ver p. 55, Baron & Banaji, 2006).

IAT Personalizado

Esta variante do IAT foi proposta por Olson e Fazio (2004), na tentativa de contornar a possibilidade dos efeitos IAT reflectirem atitudes ou crenças culturalmente partilhadas, e não as atitudes de cada participante. Com este objectivo em mente, propuseram a alteração dos típicos rótulos do par de categorias-atributo (e.g., Bom vs. Mau) para rótulos que representam julgamentos do próprio participante (e.g., Eu gosto vs. Não gosto).

Todas estas medidas alternativas têm enriquecido e complementado a literatura que se tem focado nos processos subjacentes ao efeito IAT. Mas para além da utilização destas variações do IAT, também o foco sobre os efeitos de variáveis que se evidenciou moderarem o efeito IAT, tem contribuído para o aperfeiçoamento dos delineamentos experimentais e cálculo dos efeitos.

Variáveis moderadoras

Desde o seu desenvolvimento, o IAT tem sido refinado com base em estudos que se focaram na influência que algumas variáveis exercem nos seus efeitos. A tabela abaixo apresenta algumas das mais relevantes.

Ordem dos blocos críticos	O desempenho na primeira tarefa combinada interfere com o da segunda o que conduz ao enviesamento na interpretação das associações da primeira tarefa como mais fortes. A sua ordem das tarefas combinadas compatíveis e incompatíveis deve portanto ser contrabalançada. Em alternativa pode-se aumentar o número de ensaios do bloco que precede a segunda tarefa combinada (ver Nosek et al., 2005).
Fluência cognitiva	Participantes mais lentos na realização de um IAT tendem a produzir efeitos de maior magnitude do que participantes mais rápidos, contudo, a aplicação do algoritmo D aperfeiçoado reduz esta influência à não significância (Cai, Sriram, Greenwald, & McFarland, 2004).
Idade	Participantes mais velhos tendem a obter maiores efeitos IAT do que os mais novos (Greenwald & Nosek, 2001; Hummert, Garstka, O'Brien, Greenwald, & Mellott, 2002). Esta variável relaciona-se com a fluência cognitiva, sendo a sua influência igualmente minimizada pelo algoritmo D aperfeiçoado.
Exposição prévia ao IAT	A magnitude dos efeitos IAT tende a baixar com administrações repetidas (Greenwald & Nosek, 2001; Greenwald et al., 2003). Apesar da aplicação do algoritmo D aperfeiçoado minimizar esta influência, sugere-se algum cuidado em delineamentos com vários IATs na mesma sessão ou em estudos longitudinais, aconselhando-se a inclusão no delineamento do estudo de um IAT de controlo (e.g., Flores-Insectos) (ver Nosek et al., 2007).
Ordem de aplicação das medidas explícitas	A aplicação de medidas explícitas antes ou depois da aplicação do IAT produz reducidos ou nenhuns efeitos (Nosek et al., 2005).

Variáveis tipicamente associadas ao paradigma

A aplicação do IAT, enquanto medida implícita, é geralmente acompanhada de medidas explícitas que variam consoante os constructos de interesse e tipo de IAT correspondente (e.g., estereótipos/crenças, atitudes, auto-percepção).

Assim, em casos onde se pretenda aceder a atitudes ou auto-percepções, é comum utilizar medidas explícitas com uma componente avaliativa (e.g., preferência, valência). Por outro lado, quando se

pretende aceder a crenças ou estereótipos, podem utilizar-se medidas explícitas já existentes (e.g., Escala Moderna de Racismo) ou que, regra geral, acedam a associações entre conceitos.

Com o objectivo de estudar os padrões correlacionais entre o IAT e várias medidas explícitas, uma meta-análise de Hofman, Gawronski, Gschwendner, Le e Schmitt (2005) reportou uma maior consistência entre o IAT e medidas explícitas como *ratings* de adjectivos, *feeling thermometers* e diferenciais semânticos, comparativamente a questionários de escalas mais complexos. Assim sendo, torna-se recomendável que os indicadores explícitos sejam o mais consistente possível com os indicadores implícitos, quanto ao conteúdo e especificidade (Ajzen & Fishbein, 1977).

Alguns exemplos da aplicação do paradigma

A versatilidade e facilidade de aplicação inerentes ao IAT permitiram a sua disseminação por vários contextos de investigação. Deste modo, para além da sua aplicação inicial no estudo de crenças e atitudes relativamente a exogrupos, tem-se verificado a sua utilidade no estudo da percepção das marcas pelos consumidores, auto-avaliações de personalidade, atitudes e crenças relativamente a drogas, política, género, orientação sexual e relações interpessoais (ver anexo da meta-análise de Greenwald, Poehlman, Uhlmann, & Banaji, 2009). Tal revela a utilidade no IAT na tentativa de compreender quaisquer padrões atitudinais ou de crenças que impliquem a comparação entre dois pares de conceitos-alvo.

Discriminação em contexto de recrutamento e selecção

Na tentativa de compreender se associações activadas automaticamente estariam na base de comportamentos discriminatórios por parte de profissionais de recrutamento e selecção, Rooth (2010) conduziu dois estudos de campo onde comparou a probabilidade dos recrutadores convocarem para entrevista candidatos com nomes árabes/muçulmanos ou suecos, com as suas atitudes e crenças sobre desempenho, explícitas e implícitas, face a ambas as categorias sociais. Para tal, utilizaram como medidas explícitas um *feeling thermometer*, uma escala de preferências de selecção de pessoal e uma escala de estereótipos relativos à productividade de cada grupo social, e como medida atitudinal implícita, um IAT de atitudes/estereótipos com as categorias Árabes/Muçulmanos *versus* Suecos e Alta productividade *versus* Baixa productividade. Os resultados demonstraram uma correlação negativa forte e significativa entre as atitudes implícitas e a taxa de convocações para entrevista para a categoria dos árabes/muçulmanos, que não se observou para a categoria dos suecos. Ou seja, quanto mais negativa a atitude face à productividade dos árabes/muçulmanos, menor a taxa de convocação de membros desse grupo. Os resultados deste estudo sugerem que a discriminação implícita age de forma distinta da explícita, e pode ter influência em processos de recrutamento e selecção.

Preferências implícitas e explícitas de marca e relação marca-consumidor

Para testar a eficácia do IAT enquanto medida de atitudes face a marcas, um estudo de Brunel, Tietje e Greenwald (2004) avaliou a atitude face a duas marcas e a relação destas com o auto-conceito dos consumidores, tanto a nível implícito, por recurso a dois IATs (um atitudinal e outro de auto-conceito), como a nível explícito, através de medidas que acederam à preferência de marca, frequência de utilização e propriedade de itens de cada marca. Para o efeito, seleccionaram um tópico de baixa sensibilidade social, de forma a que as respostas explícita e implícita não divergissem em consequência

da desejabilidade social (*Apple/Macintosh vs. PC-type/IBM*). Enquanto no IAT atitudinal as categorias-alvo selecionadas foram combinadas com as categorias-atributo Agradável *versus* Desagradável, no IAT de auto-conceito as categorias-atributo foram Eu *versus* Outros. Os resultados demonstraram que os dois IATs se correlacionavam com as medidas explícitas. Os IATs permitiram diferenciar os consumidores com base na preferência, posse e utilização de uma marca relativamente à outra. Adicionalmente, observou-se ainda que os efeitos IAT relativos à relação marca-consumidor (IAT de auto-conceito), foram superiores para os adeptos do Mac, sugerindo que a posse de um *Mac* está muito mais relacionada com a identidade implícita de cada um, o que não parece acontecer com os adeptos do *PC*. Este estudo demonstra a validade do IAT enquanto medida de atitudes implícitas face às marcas e da relação marca-consumidor, e sustenta a sensibilidade desta medida ao nível individual (mas para uma crítica à interpretação de efeitos IAT ao nível individual ver Fiedler, Messner, & Bluemke, 2006).

A explicação teórica

Nesta secção focam-se os modelos teóricos que sustentam as suposições que estão na base do IAT, e que fornecem uma explicação para os seus efeitos.

Princípio da dispersão da activação

A premissa na qual o IAT se baseia alicerça-se nos pressupostos da teoria ACT (*Adaptive Control of Thought*) desenvolvida por John R. Anderson (para uma revisão ver Anderson, 1983), mais especificamente, no princípio da dispersão da activação. Esta teoria propõe uma arquitectura da memória onde o conhecimento é representado sob a forma de nódulos associados entre si através de ligações numa rede semântica. Segundo o princípio da dispersão da activação, a activação de elementos de informação (ou fontes de activação) na memória de trabalho, dispersa-se através das ligações entre os nódulos da rede até à memória de longo-prazo. Os níveis de activação na rede reflectem o grau de associação entre as representações existentes e as fontes de activação, variando a velocidade de processamento da informação em função do nível de activação. Deste modo, o estímulo “traça” agirá como fonte de activação e irá mais rapidamente activar a representação de “insectos” do que a de “flores”, o que poderá explicar a maior rapidez de categorização do estímulo na categoria no IAT. Caso o estímulo “traça” esteja mais fortemente associado a representações de valência negativa, num bloco crítico congruente com as categorias Insectos e Negativo codificadas para a mesma resposta, a tarefa de categorizar “traça” e “ódio” utilizando a mesma resposta irá ser facilitada em consequência da sua proximidade na rede.

Explicações alternativas do efeito IAT

Embora o princípio da dispersão da activação de Anderson (1983) dê suporte ao IAT, resistem dúvidas relativamente ao mecanismo que produz os efeitos IAT e ao que realmente poderá estar a medir. Na procura de uma explicação teórica compreensiva para os mecanismos subjacentes aos efeitos IAT, surgiram, ao longo do tempo, diversos estudos que empreenderam no seu desenvolvimento. Entre as diferentes propostas alternativas, o que parece não variar entre as várias explicações alternativas dos mecanismos subjacentes ao efeito IAT, é o foco no grau de similaridade entre aspectos das categorias-alvo e de atributo. Em suma, parece que quanto maior é a similaridade entre os dois pares de categorias,

mais similares parecem também ser os padrões de activação a que dão origem, o que resulta em respostas facilitadas quando estes são codificados para a mesma resposta num bloco congruente. Embora as similaridades possam ter múltiplas origens, pode ser que a partilha destas entre conceitos forneça a base para as associações entre os mesmos (Fazio & Olson, 2003). De seguida são descritas algumas das contribuições mais relevantes para a compreensão dos efeitos IAT.

A hipótese da categorização com base em aspectos estruturalmente relevantes ou irrelevantes. Numa observação do procedimento do IAT, De Hower (2001) notou que o aspecto estruturalmente relevante de qualquer estímulo de uma categoria-alvo (e.g., rosa) é a sua pertença a uma das categorias-alvo (e.g., rosa é uma flor, não um insecto). Porém, note-se também que o IAT é sensível a aspectos estruturalmente irrelevantes como é o caso da valência dos estímulos (e.g., rosa tem valência positiva). Se tal acontecer, torna-se difícil garantir que nas tarefas combinadas congruentes os estímulos estejam a ser classificados com base na categoria a que pertencem, ou apenas com base na sua valência, dado que a resposta é a mesma.

Para testar se o desempenho num IAT era função da valência das categorias ou dos exemplares de cada uma, concebeu um IAT com as categorias “Britânicos *versus* Estrangeiros” e “Positivo *versus* Negativo”, manipulando a valência dos estímulos das categorias-alvo por forma a incluir tantos estímulos positivos (e.g., Princesa Diana para Britânicos; Albert Einstein para Estrangeiros) quanto negativos (e.g., Jack o Estripador; Adolf Hitler), que aplicou a uma amostra de participantes britânicos. Os resultados demonstraram que o desempenho apenas foi afectado quando as categorias Britânicos e Positivo eram mapeadas para a mesma ou diferente resposta comportamental, não se observando qualquer efeito relativamente à valência dos estímulos individuais como inicialmente hipotetizado. Deste modo, conclui que o desempenho num IAT é mais ditado pelas instruções da tarefa do que por possíveis classificações com base em aspectos estruturalmente irrelevantes como a valência dos estímulos.

A estratégia de evitamento da alternância entre discriminações alvo-atributo. Mierke e Klauer (2003) focaram-se na possibilidade dos participantes ignorarem a instrução das tarefas combinadas de que devem alternar entre a tarefa de atribuir estímulos à categoria-alvo *versus* à categoria-atributo, pela razão de a atribuição com base somente num dos pares de categorias (alvo ou atributo) parecer ser suficiente para responder rápida e correctamente à tarefa. Assim, os participantes evitam custos associados à alternância entre a discriminação de estímulos com base em cada par de categorias, o que permite uma eficiência maior nos blocos congruentes. No entanto, nos blocos incongruentes, onde a alternância entre categorias se torna inescapável, a estratégia acaba por se revelar contra-producente, dando origem a TR mais elevados. Ao compararem TR entre ensaios que solicitavam a alternância da discriminação entre categoria-alvo e atributo, e ensaios que solicitavam o mesmo tipo de discriminação, repararam que alternar entre discriminações alvo-atributo produzia TR mais elevados nos blocos incongruentes do que nos blocos congruentes do IAT. Estas evidências sugerem que nem as assimetrias de saliência, nem associações entre conceitos previamente existentes, sejam pré-condições necessárias para a obtenção de efeitos IAT.

O modelo da assimetria figura-fundo. Rothermund e Wentura (2001, 2004) propõem como explicação subjacente ao efeito IAT um modelo de figura-fundo, que postula uma assimetria na saliência entre as categorias-alvo e as de atributo. Segundo este modelo, a saliência de uma categoria-alvo (e.g., Negros) e respectiva categoria-atributo (e.g., Negativo) funcionará como uma “figura” num “fundo” representado pelas categorias opostas não-salientes (e.g., Brancos e Positivo). Assim, quando no bloco congruente as categorias salientes se encontram mapeadas para uma tecla, e as não-salientes para a outra, os participantes poderão basear a sua discriminação com base na assimetria de saliência por si só. Por outro lado, no bloco incongruente, uma incompatibilidade entre as categorias salientes e não-

salientes ocorre. Os autores dissociaram efeitos de saliência assimétrica de efeitos de associações entre conceitos-alvo e atributos, demonstrando como os efeitos IAT podem emergir apenas com base na assimetria de saliência. Porém, tal não descarta a hipótese das diferenças entre as forças de associação poderem também estar na base dos efeitos IAT, ou de que as assimetrias de saliência possam ser resultado dessas mesmas diferenças. Como tal, a categoria Insectos pode ser a categoria saliente de um IAT Flores-Insectos devido aos insectos estarem mais associados a valência negativa.

Em defesa dos efeitos IAT se poderem basear mais nas características nominais das categorias do que na assimetria da sua saliência, Greenwald, Nosek, Banaji e Klauer (2005) testaram esta interpretação de recodificação estratégica de Rothermund e Wentura, concluindo que a assimetria de saliência só distorce os efeitos IAT em IATs concebidos para maximizar assimetrias de valência e minimizar associações entre as categorias.

Críticas ao IAT

Embora o IAT se tenha tornado uma das medidas implícitas mais populares nos últimos anos, a ausência de um modelo teórico robusto que permita explicar os mecanismos que conduzem aos efeitos IAT colocou-o sob a atenta vigilância da comunidade científica. As explicações alternativas supramencionadas ilustram como a típica interpretação dos resultados de um IAT pode ser precipitada, principalmente quando deles se inferem atitudes implícitas.

As evidências encontradas por Mierke e Klauer (2003) impedem não só que modelos explicativos alternativos como o da assimetria figura-fundo (Rothermund & Wentura, 2001, 2004) se estabeleçam como explicação plausível dos mecanismos subjacentes ao efeito IAT, como ainda coloca em causa a prévia existência das associações, que sejam medidas pelo IAT, fora do contexto da sua aplicação.

Outra das críticas dirigidas ao IAT prende-se com a sua utilização enquanto ferramenta de diagnóstico individual, ou por outras palavras, considerar que o índice obtido num IAT por um indivíduo seja diagnóstico das suas atitudes. Como demonstram Fiedler et al. (2006), a probabilidade de obter um D positivo num IAT é maior do que a probabilidade de a atitude realmente existir no participante, o que coloca um sério obstáculo à inferência diagnóstica da atitude com base no resultado do IAT, visto que este pode representar um falso positivo. Por outro lado, caso o IAT seja aplicado em condições onde se tenha comprovado (ou mesmo promovido) a existência da atitude, deduz-se que a probabilidade de obter um falso positivo diminua consideravelmente (Fiedler et al., 2006). Deste argumento depreende-se uma maior utilidade do IAT enquanto teste confirmatório do que enquanto ferramenta de diagnóstico. Está ainda implícito nesta crítica que os efeitos IAT podem reflectir outras causas para além de uma atitude (ou crença, estereótipo), uma observação consistente com as conclusões de Mierke e Klauer (2003).

As críticas aqui apresentadas tornam-se importantes aquando da interpretação dos resultados obtidos em estudos que utilizem o IAT, ao mesmo tempo que contribuem para o contínuo refinamento do paradigma associado ao IAT.

Sumário

Desde o seu aparecimento, o Teste de Associação Implícita (IAT) tem sido utilizado como medida implícita de eleição para aceder aos mais variados constructos tais como atitudes, crenças ou

estereótipos, em diversos contextos de investigação (e.g., clínica, organizacional, política) para além da cognição social.

A sua versatilidade e facilidade de aplicação rapidamente contribuíram para que se afirmasse como um novo paradigma, imprescindível na compreensão das relações associativas entre os mais variados conceitos e representações cognitivas.

O presente artigo descreve em pormenor os aspectos metodológicos associados ao IAT, informando ainda sobre as variáveis que moderam os efeitos IAT e que devem ser consideradas no planeamento dos estudos. São referidas algumas das explicações teóricas para os mecanismos subjacentes ao efeito IAT que têm sido propostas na literatura, e ainda resumidas as principais críticas de que o IAT tem sido alvo.

Artigos a ler

- Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, 22(3), 261-295.
- Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Journal of Personality and Social Psychology*, 102, 4-27.
- Fiedler, K., Messner, C., & Bluemke, M. (2006). Unresolved problems with the “I”, the “A”, and the “T”. *European Review of Social Psychology*, 17, 74-147.
- 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.
- 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.

Referências

- Ajzen, I. T., & Fishbein, M. (1977). Attitude – Behaviour relations: A theoretical analysis and review of empirical research. *Psychological Bulletin*, 84, 888-918.
- Anderson, J. R. (1983). A spreading activation theory of memory. *Journal of Verbal Learning and Verbal Behavior*, 22(3), 261-295.
- Baron, A. S., & Banaji, M. R. (2006). The development of implicit attitudes: Evidence of race evaluations from ages 6, 10 & adulthood. *Psychological Science*, 17, 53-58.
- Brunel, F. F., Tietje, B. C., & Greenwald, A. G. (2004). Is the Implicit Association Test a valid and valuable measure of implicit consumer social cognition? *Journal of Consumer Psychology*, 14, 385-404.
- Cai, H., Sriram, N., Greenwald, A. G., & McFarland, S. G. (2004). The Implicit Association Test’s D Measure can Minimize a Cognitive Skill Confound: Comment on McFarland and Crouch (2002). *Social Cognition*, 22(6), 673-684.

- De Houwer, J. (2001). A structural and process analysis of the Implicit Association Test. *Journal of Experimental Social Psychology*, 37, 443-451.
- De Houwer, J., & De Bruycker, E. (2007). The implicit association test outperforms the extrinsic affective Simon task as an implicit measure of inter-individual differences in attitudes. *British Journal of Social Psychology*, 46, 401-421.
- Fazio, R. H., & Olson, M. A. (2003). Implicit measures in social cognition research: Their meaning and use. *Annual Review of Psychology*, 54, 297-327.
- Fiedler, K., Messner, C., & Bluemke, M. (2006). Unresolved problems with the "I", the "A", and the "T". *European Review of Social Psychology*, 17, 74-147.
- Foroni, F., & Bel-Bahar, T. (2010). Picture-IAT versus Word-IAT: level of stimulus representation influences on the IAT. *European Journal of Social Psychology*, 40(2), 321-337. doi: 10.1002/ejsp.626
- Greenwald, A. G. (2005). The Multifactor Trait IAT. Unpublished manuscript. University of Washington, Seattle.
- Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. *Journal of Personality and Social Psychology*, 102, 4-27.
- 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.
- Greenwald, A. G., Nosek, B. A., Banaji, M. R., & Klauer, K. C. (2005). Validity of the salience asymmetry interpretation of the IAT: Comment on Rothermund and Wentura (2004). *Journal of Experimental Psychology: General*, 134, 420-425.
- 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.
- Greenwald, A. G., & Nosek, B. A. (2001). Health of the Implicit Association Test at age 3. *Zeitschrift für Experimentelle Psychologie*, 48(2), 85-93.
- Hofmann, W., Gawronski, B., Gschwendner, T., Le, H., & Schmitt, M. (2005a). A metaanalysis on the correlation between the Implicit Association Test and explicit self-report measures. *Personality and Social Psychology Bulletin*, 31, 1369-1385.
- Hummert, M. L., Garstka, T. A., O'Brien, L. T., Greenwald, A. G., & Mellott, D. S. (2002). Using the implicit association test to measure age differences in implicit social cognitions. *Psychology and Aging*, 17(3), 482-495.
- Karpinski, A., & Steinman, R. B. (2006). The single category implicit association test (SC-IAT) as a measure of implicit social cognition. *Journal of Personality and Social Psychology*, 91, 16-32.
- Lane, K. A., Banaji, M. R., Nosek, B. A., & Greenwald, A. G. (2007). Understanding and using the Implicit Association Test: IV. What we know (so far) (pp. 59-102). In B. Wittenbrink & N. S. Schwarz (Eds.), *Implicit measures of attitudes: Procedures and controversies*. New York: Guilford Press.
- Lemm, K. M., Lane, K. A., Sattler, D. N., Khan, S., & Nosek, B. A. (2008). Assessing implicit attitudes with a paper-format Implicit Association Test. In T. G. Morrison & M. A. Morrison (Eds.), *The psychology of modern prejudice*. Hauppauge, NY: Nova Science Publishers.
- Mierke, J., & Klauer, K. C. (2003). Method-specific variance in the Implicit Association Test. *Journal of Personality and Social Psychology*, 85, 1180-1192.

- Mitchell, J. A., Nosek, B. A., & Banaji, M. R. (2003). Contextual variations in implicit evaluation. *Journal of Experimental Psychology*, *132*(3), 455-469.
- Nosek, B. A., & Banaji, M. R. (2001). The go/no-go association task. *Social Cognition*, *19*(6), 625-666.
- Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2005). Understanding and using the Implicit Association Test: II. Method variables and construct validity. *Personality and Social Psychology Bulletin*, *31*, 166-180.
- Olson, M. A., & Fazio, R. H. (2004). Reducing the influence of extra-personal associations on the Implicit Association Test: Personalizing the IAT. *Journal of Personality and Social Psychology*, *86*, 653-667.
- Rooth, D.-O. (2010). Automatic associations and discrimination in hiring: Real world evidence. *Labour Economics*, *17*, 523-534.
- Rothermund, K., & Wentura, D. (2001). Figure-ground asymmetries in the Implicit Association Test (IAT). *Zeitschrift fuer Experimentelle Psychologie*, *48*(2), 94-106.
- Rothermund, K., & Wentura, D. (2004). Underlying processes in the Implicit Association Test (IAT): Dissociating salience from associations. *Journal of Experimental Psychology: General*, *133*(2), 139-165.
- Sriram, N., & Greenwald, A. G. (2009). The Brief Implicit Association Test. *Experimental Psychology*, *56*, 283-294.

