Commentary

IAT(潜在連合テスト)を用いて「無意識の偏見」を測る

アンコンシャス・バイアス(無意識の偏見)測定を社会調査に取り込む

潮村公弘
フェリス女学院大学文学部 教授

1 はじめに

アンコンシャス・バイアス(Unconscious Bias)という概念は,ここ数年で,日本国内において急速に認知度が高まってきた概念であり,最も取り上げられることが多いのは組織での人材マネジメントの場ではないかと思われる。ほかにも,男女共同参画や伝統的なステレオタイプ論等においても幅広く言及されてきている。

このアンコンシャス・バイアス(無意識の偏見/無意識のバイアスとも訳される)は,心理科学領域において,「潜在的連合」あるいは「潜在態度」という心理学用語で研究が進められてきた概念を指し,自分が意識することができない(あるいは,少なくとも明確に意識することができない)概念間の連合(結びつき)のことである。

アンコンシャス・バイアスは,意識することができないという性質上,社会調査における通常の測定方法では測定ができない。このアンコンシャス・バイアス(すなわち潜在的な連合)を測定するための心理科学技法は複数開発がされてきているが,その中で,デファクト・スタンダードと言える測定技法がImplicit Association Test(IAT:読み方はアイ・エイ・ティー)である。

かつてThurstone(1928)は,「Attitudes can be measured」という心理学における金字塔的な論文を著した。それまでは,態度などというものは測定できないと考えられていたのである。IAT開発の中核的研究者であるBanajiはその約70年後に「Implicit attitudes can be measured」(Banaji, 2001)という論考を発表し,潜在連合テスト(IAT)によって潜在的態度(Implicit Attitude)は科学的に測定できるようになったことを主張した。

筆者は,心理学系の研究組織と社会学系(主に社会調査系)の研究組織との両者に所属する機会を有した経緯から,IATを用いて測定可能な潜在態度と,一般的な社会調査で測定される顕在態度との統合的理解を自身の研究目標に据えてきたという経緯を持つ。一見,大きくかけ離れているように思える潜在態度と顕在態度の両測度をともに取り上げることで,まったく新しい研究テーマや研究領域を切り開いていけるものと考えている。

2021年7月には,筆者自身によるIATの体験サイト(各自でIATを実際に体験し,回答者自身の潜在バイアスをあらわすIATスコアーのフィードバックを得られるサイト)及び,研究用データ収集サイトのローンチを行った。今後,さらに幅広くIATを採用した研究プラットフォームの提供が可能となるように,さらなる開発を進めている(文末のサイト紹介を参照)。

2 潜在連合テスト（IAT）とは

2-1節：潜在連合とは

心理科学における基礎概念に，「連想価」という概念がある。「連想価」とは単語・概念間の連想の程度を表す基礎的データで，元々心理学の実験研究での刺激（刺激語）選びのために整理されているものである。なお，刺激（刺激語）とは，実験課題において回答者が反応すべき対象を指す。このような「連想価表」が存在していることは，日本語使用者において，単語・概念間の連想価がかなりの程度，共有されていることをあらわしている。これは，社会的ステレオタイプの共有性の高さをあらわしていることの反映でもある。

この単語・概念間の「連想価」とは，個々の言葉や概念が，私たちの記憶構造（知識構造とも言う）において，それぞれに特有の強度で結び付いている様相を反映したものと言える。Collins & Loftus（1975）は，古典的研究において，記憶構造内での諸概念間の関係性を示す模式図を示している（図1）。図1に示されているように，「赤（RED）」に対して，私たちは容易に素早く「火（FIRE）」や「リンゴ（APPLES）」を連想するが，「梨（PEARS）」は「リンゴ」を介することでのみ連想が生じる。

このように，刺激を見たり聞いたり想起する（思い出す）ことで概念が活性化され，その活性化が「知識構造」に基づいて伝播（拡大）していくという考え方を，活性化拡散理論と呼ぶ。人の意識や判断は，このような記憶内の知識構造の反映であると考えられる。このことを，神経ネットワーク内での電気信号の伝達として説明する立場もあるが，ここではそのような神経科学的な背景には深入りせず，単語や概念には，「連合（潜在的な連合）」があり，その活性化は無意識的に生起する，ということを確認しておきたい。

このような潜在的な連合をどのようにして測定できるかについて，心理科学（とくに認知系の社会心理学研究）では長い研究の歴史がある。

例えば，非意識的な刺激が人の評価・感情・行動に及ぼす影響を実証したBarghによる一連の研究（Bargh, 1997など）や，現代アメリカ人では心理尺度で測定された偏見が弱い群の人たちにも，偏見が強い群の人たちと同じような偏見づけられた知識構造が存在していることを示したDevineの一連の研究（Devine, 1989など）などが代表的な研究であり，これらの知見の上に，Greenwaldら（1998）のIATの開発が位置している。

2-2節：潜在連合テスト（IAT）の原理と実際

潜在連合テスト（IAT）は，Greenwaldら（1998）で発表された心理科学における測定技法である。

IAT課題について説明をするさいの最もシンプルな課題例が「花」「虫」を「良い」「悪い」に分類する課題である。〈速く〉〈正確に〉分類できる組み合わせは，「花」と「良い」をペアーに，「虫」と「悪い」をペアーにする分類課題であると言える。この課題で測定できるものは，多く

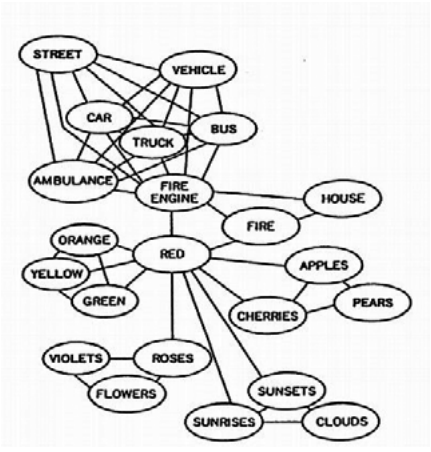

図1　記憶構造内での概念間の関係性を表す模式図
（Collins & Loftus,1975より）

の人が有している「花」と「虫」に対する潜在連合であり，潜在態度であると位置づけられる。

実際の潜在連合テスト（IAT）では，典型的には2つの概念（例：「自分」「他者」）と2つの属性（例：「良い」「悪い」）を取り上げ，図2と図3のような課題遂行画面をディスプレイ上に構成する。ここでは，IAT課題として自己に対する「潜在的自尊心」IATを取り上げる。

図2では，ディスプレイ中央に呈示される刺激（単語や画像などが用いられる）が「自己」か「良い」のいずれかに属する刺激である場合には左のキーをできるだけ速く押し，「他者」か「悪い」に属する刺激の場合には右のキーを押す，という反応が求められる。一般に，「自己」は「良い」ものと捉えられているので，この課題は「一致課題（congruent block）」と呼ばれる。

一方，図3では，ディスプレイ中央に呈示される刺激が「自己」か「悪い」のいずれかに属する刺激である場合には左のキーをできるだけ速く押し，「他者」か「良い」に属する刺激の場合には右のキーを押す，という反応が求められる。この図3で示されている課題は，「不一致課題（incongruent block）」と呼ばれる。

図2で示されている一致課題では，通常，反応が容易であるため，速く反応することができ，また，押下すべきキーとは反対側のキーを押下してしまう「エラー（誤答）」は少ない。なお，回答者の反応が「エラー（誤答）」である場合には，刺激の下に「×」がフィードバックされる。図3に示されている不一致課題は，通常，難易度が高く，反応するまでの時間が長くかかり，エラーも多くなる。

単純に記せば，「不一致課題での反応時間の平均」から「一致課題での反応時間の平均」を引いた差異が，回答者の潜在（的）連合の強さをあらわす指標である。そして，この指標の値が，回答者自身が有しているアンコンシャス・バイアス（無意識の偏見）のレベルを示す（なお，実際の研究ではCohenのdスコアーを用いることが推奨されている。IAT課題に特化したdスコアー算出式はGreenwaldら（2003）において提唱されている）。

2-3節：潜在連合テストの信頼性と妥当性

先にも記したように，この潜在連合テスト（IAT）は，潜在意識や潜在連合を測定するためのデファクト・スタンダードとなっている。それゆえ，測定技法（測定指標）としての信頼性・妥当性の問題をはじめ，測度としての適切さの検証も指標の開発当初から活発に行われている。ここでは，この点について簡単に紹介していきたい。

まず基本的なこととして，潜在連合テスト（IAT）のスコアー（IAT効果量）は，課題を構成する以下の諸要因には影響を受けないことが実証されている（Greenwald & Nosek, 2001）。具体的には「概念／属性を左右どちらのキーに割り当てるか」「カテゴリーを構成する刺激語の数」「試行間間隔＝回答者の反応から次の刺激呈示までの間隔」「右利きか，左利きか，両

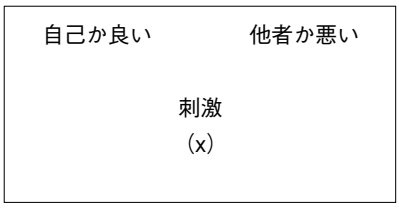

図2　一致（congruent）課題

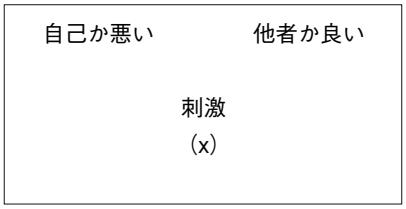

図3　不一致（incongruent）課題

手利きか」「概念を表すための刺激語の熟知性（familiarity：よく知っている言葉かどうか）」などである。

次に，信頼性・妥当性についても多くの検討がなされている。潜在連合テスト(IAT)は新しい測度（指標）であるので，この点は，クリティカルな問題であると言える。信頼性については，「再検査信頼性」や「内的一貫性」の点から検証が行われている。また，妥当性についても「内的妥当性」や「予測的妥当性」等の点から検証が行われており，潜在連合テスト（IAT）は，信頼性・妥当性の観点からも潜在意識指標のデファクト・スタンダードとして位置づけられていると言える。

さらには，潜在連合テスト（IAT）でのスコアー（IAT効果量）は，回答者自身の意向や回答者の回答方略の影響を受けないこと，すなわち，回答者が自分の都合の良いような測定値を導きたいと考えてもそれはできないことも実証されている。回答者が自身の意向に沿った測定結果を示すことができたとしたら，それは意識的な顕在測度と変わらないものになってしまう。潜在連合を測定するIATでは回答者が回答を偽ることはできないのである。なお，信頼性・妥当性の詳細については，拙著・潮村（2016）を参照いただきたい。

3 潜在連合テストの測定バリエーションと社会調査への活用

3-1節：潜在連合テストの測定上のバリエーション

測定上のバリエーションとして，「測定方法（測定環境）のバリエーション」と，「測定技法のバリエーション」があり，IATはこのどちらにおいても多様な選択肢が存在しており，社会調査の実施環境や研究トピックに応じて多様なバリエーションからの選択が可能である。

前者のバリエーションとしては，大別して，〈パソコン（PC）版IAT〉〈紙筆版IAT〉〈Website版IAT〉があり，さらにはタブレット端末やスマートフォンでの利用も行われてきている。

〈パソコン（PC）版IAT〉 スタンダードな測定方法は〈パソコン（PC）版IAT〉であり，本稿で前述してきた測定方法はパソコン（PC）版IATに依拠している。パソコン（PC）版IATを実施するためのプログラムはInquisitというソフトウェアを利用することがスタンダードになっているが，他のプログラミング言語を用いても全く問題はない。

IAT開発の中心的研究者であるGreenwaldやNosekのウェブサイトには，実施をサポートしてくれる様々な資料素材がアップされている。加えて，Inquisitサイトでも，サンプルプログラムの提供をはじめ交流サイトなども開設されている。

〈紙筆版IAT〉 パソコン（PC）を用いた測定が測度の厳密性の点で望ましいと言えるが，パソコンが必須であることは，とくに集合自記式の調査環境ではハードルが高い。そのため，IATには開発初期の段階から紙とペンを用いて回答する紙筆版が開発され，利用されてきた。

具体的な測定方法については，紙面の制約から潮村（2015）を参照していただきたい。当該論文を参照するのみで，紙筆版IATを自身で遂行できるように構成されているとともに，論文掲載誌である『対人社会心理学研究』のウェブサイト上からは，紙筆版IATを実際に遂行するための資料ファイル一式（電子ファイル）が提供されている。この電子ファイルをダウンロードして文言を置換することによって，紙筆版IATの実施と集計（得点化）が実現できるように構成されている（説明資料：A4用紙3枚，回答用紙：A4用紙8枚，紙筆版IATの所要時間は12分程度の構成となっている）。

〈Website版IAT〉 Website版IATは，特定のWebsiteにアクセスすることで，回答者自身のパソコン（PC）上で，IATを遂行できるように開発されたシステムである。現行では，主に二

つの方法があり，一つは，Inquisit Webというソフトウェアを購入して実施する方法である。もう一つは，IAT専用のウェブサイトを開発する方法であり，この場合は本格的なアプリケーション開発が必要となる。

〈さらなる応用型IAT：タブレット版／スマートフォン版〉 上記に加えて，さらなる応用型のIATの測定方法（測定環境）として，タブレット端末版やスマートフォン版も開発され，利用されている。

一方，後者の「測定技法のバリエーション」としては，標準的なIATに加えて，キーを押すか押さないかの判断を求めるGNAT，複数の概念ではなく単一の概念に対する潜在連合を測定するための広義のSingle Category IAT（SC-IAT），回答者個々人の名前や所属集団等の個別記述的情報を課題に取り込んだidiographic IATなどの多様なバリエーションがある。

3-2節：社会調査や研修プログラムへのIAT活用のバリエーション

ここでは，一般的なスタイルの社会調査や，e-ラーニング，オンライン研修等での潜在連合テスト（IAT）活用のバリエーションについて紹介していきたい。

教育現場等での集合自記式調査方式は回収率も高く効率的な調査手法であり，潜在連合テスト（IAT）を活用するには，「紙筆版IAT」をはじめ，実施環境によっては「タブレット版IAT」「スマートフォン版IAT」の活用が有効であると思われる。実施環境に応じて，どの方法がベストな方法かを考えて選択することができる。

次に，近年増加してきた調査方法であるオンライン調査（Web調査）においては，オンラインで回答している状態のままで「Website版IAT」へと導き入れることが望ましいであろう。この場合，先にも述べたように，自身でInquisit Webを利用して「Website版IAT」を構築することが可能である（ただし金銭的コスト面でのハードルは少し高い）。

e-ラーニングやオンライン研修において，「Website版IAT」を活用することで，回答者のアンコンシャス・バイアス（＝潜在連合）の測定と自身のバイアスへの気づきを与えることが可能であり，この方法は教育効果や研修効果の向上に有用であると考えられる。厳密なIATスコアーを回答者にフィードバックする方法は専門性が高くなるが，簡易的なフィードバックでも回答者が「気づき（awareness）」を得る効果はある。そのため，ハーバード大学版IAT（Project Implicit）の日本サイト（https://implicit.harvard.edu/implicit/japan/）が用意しているIAT課題を経験することでもかなりの程度の教育効果・啓蒙効果が期待できる。回答結果を自身でコントールすることはできないものの，遂行していく過程で自分のアンコンシャス・バイアスがどの方向性のものか（アンコンシャス・バイアスがあるのか，また強いかどうか）はおおむね覚知することができることから，厳密なIATスコアーが直接的にフィードバックされることがなくとも，実際に課題を体験することにより，覚知効果を享受することが可能となる。

なお，日本国内での先駆的なプログラムとして，チェンジウェーブ社のダイバーシティ（多様性の尊重）に関するe-ラーニングツール「ANGLE」や，日本ブラインドサッカー協会による，人が無意識に持つ偏見を可視化するプログラム「UB-Finder」を紹介しておきたい（文末のサイト紹介を参照）。

4 潜在連合テストを活用した潜在測定が切り開く新領域

ここでは潜在連合テスト（IAT）が切り開く新たな研究テーマについて整理していきたい。

（1）社会的望ましさ（自己呈示動機）の影響を排除可能

まず，潜在連合テスト（IAT）の利点の中核は，

回答時における「社会的望ましさ」の影響を排除できる点である。言い換えれば, 回答者が自身をより良く見せたいという自己呈示動機の影響を除外できるというメリットが挙げられる。既存の顕在指標においても社会的望ましさへの対処は様々に検討されてきたが, 決定打とはなっていないと思われる。

(2) 回答時の係留効果を回避した測定が可能

次に, 顕在指標では回答時に不可避的に生じてしまう係留効果（アンカリング効果／anchoring effectとも呼ばれる）を回避できることを指摘できる。顕在指標では, 以前の回答を自身が覚えている（あるいは覚えていないとしても想像する）ことで, 以前の測定が後続の測定に影響を与えてしまうことが避けられない。しかし, IAT課題では, 以前の測定値を覚えていようがなかろうが, 回答を意識的にコントロールできないので, 係留効果が生じないことになる。

係留効果を回避できることは, 測定論上, 重要な論点であるとともに, さらに一歩進めて, 意識や態度の, 動的で可変的な側面の敏感な測定が可能になる。Shiomura（2002）は, 目の前にいる他者から受容的ではない否定的態度を示された前と後であっても, 状態自尊心尺度（特性自尊心ではなく, 状態的な自尊心を測定できるとする尺度）での測定値は回答者全体として変化を示さなかった一方で, IATで測定した潜在的自尊心の測定値は変化したことを実証している。

(3) 意識や態度の形成過程・変容過程にアプローチ可能

ある特定の意識や態度が, なぜ, どのような経緯で, 形成されたのかを扱うことは難しい。また, 意識や態度が, 何が原因で変容して行くのかを描き出すことも同じように難しいと言えるだろう。潜在連合テスト（IAT）では, 実験操作（実験研究での変数操作）への親和性が高いことから, 意識や態度の形成過程・変容過程にアプローチする様々な研究に活用されている。代表的なものに, 人工的に潜在意識・潜在態度を創り出す研究（例えば, Laneら, 2001）や, スティグマ化されたカテゴリーに属する人たちの中で高い評価を得ている人たちを繰り返し思い起こしてもらうことで回答者の潜在態度を肯定的なものに変容できることを示した研究（Dasgputa & Greenwald, 2001）, 反ステレオタイプ的な思考を繰り返し想起することで, 潜在的ステレオタイプを低減できることを実証した研究（Blairら, 2001）などがある。

(4) 自身への覚知の効果を扱うことが可能

Website上でIATを体験できるように設計されたプログラム（Harvard大学のProject Implicitサイト, Harvard大学サイト内の日本サイト, 文末のサイト紹介欄で紹介しているUB-Finderなど）では, 回答者のIATスコアー（潜在バイアスのスコアー）が明示的に数値でフィードバックされているかどうかにかかわらず, 回答者自身への覚知（気づき）の効果が生じる。もちろんパソコン（PC）版IATでも, 紙筆版IATでも, あるいは体験的な簡易版メニューとして, スクリーンやディスプレイに呈示された刺激に対してキーを押す代わりに左右の膝を軽く叩くというような体験であってもこの覚知の効果が期待され, 教育的プログラムとしての活用ができる。

(5) 非意識（無意識）という未開拓で広大な領域に取り組むことが可能に

そして何よりも, 人に関わる科学が長年にわたって追い求めてきた非意識（無意識）へのアプローチという, 広大な領域が未開拓のまま残されており, この領域に取り組めることの魅力を指摘できるだろう。

5 おわりに

現在主流である社会意識調査の回答形式（顕在意識測定）は，今後も無くなるようなことはないだろうが，それらは回答形式の一つという位置づけになっていくだろうという意見をよく聞く。また例えば，実際の投票行動や消費者行動をより良く予想できるのは，顕在意識ではなく潜在意識であるという主張もある。

既に日本国内においても，潜在連合テスト（IAT）を活用した潜在意識（潜在連合）の測定とその測定を通じた教育プログラムや研修プログラムが提供されてきており，成果を上げている。本稿の最後に，筆者が関わったプログラムを中心に文末のサイト紹介欄で具体的に紹介しておきたい。

社会調査に潜在意識測定を取り込むことは，海外においては心理学領域以外にも社会科学で既に広く行われており，日本国内の社会調査においても潜在連合テスト（IAT）の活用をはじめとした，潜在意識測定がごく普通に行われる日が遠からず訪れるものと期待している。

私たちの意識や態度，動機，そして行動を的確に理解するには，既存の顕在指標のみでは不十分であり，潜在連合テスト（IAT）に代表される潜在指標をも活用することが不可欠であることは間違いのないことと言えるであろう。そして，そのための手法は既に存在しているのである。

IATによる測定は，心理科学・社会科学がようやく手にしえた新しい測度（科学的な潜在測度）である。この潜在測度を用いることで，新しいリサーチ・クエスチョンが開拓され続けることが期待される。

サイト紹介　IATを実際に体験できるサイトや，IATについて説明・紹介されているサイト

◎〈ハーバード大学内のProject ImplicitによるIATサイト〉と，Project Implicit日本サイト
https://implicit.harvard.edu/implicit/japan/
https://implicit.harvard.edu/implicit/

◎NPO法人日本ブラインドサッカー協会開発のサイト
日本ブラインドサッカー協会2020「人が無意識に持つ偏見を可視化する」プログラム「UB-Finder」を開発～プログラムを一般・法人に公開（URL: https://www.ub-finder.org）
https://www.b-soccer.jp/news/14591-pr200731_iat-announcement

◎一般社団法人 男女共同参画学協会連絡会の情報サイト内での紹介
男女共同参画学協会連絡会「無意識のバイアス」中の「SEE BIAS and BLOCK BIAS」ページ
https://www.djrenrakukai.org/unconscious-bias/see_bias_block_bias/index.html

◎『対人社会心理学研究』誌（第15巻）のデータアクセスサイト（潮村，2015）
同誌のWebsiteから，論文で提供した，実施・遂行用の各種電子ファイルを入手することが可能。
http://syasin.hus.osaka-u.ac.jp/jjisp/no15.html

◎IATを活用した，チェンジウェーブ社プロダクト「ANGLE」（日本の人事部「HRアワード2019」組織変革・開発部門 優秀賞受賞）紹介サイト
https://changewave.co.jp/product/

◎〈Explore Implicitプロジェクト：IATデータ収集システム開発【2021年7月ローンチ】〉
潜在意識測定のためのデータ収集新システムを構築して稼働開始。〈教育目的を重視した体験的IATサイト（兼，データ収集用IATサイト）〉に加え，外部の研究者との共同研究用のIATデータ収集サイトである〈共同研究用IATサイト〉も並行して運用をスタート。なお今後，研究者向けに研究利用の自由度が高く活用しやすいデータ収集システムにアップグレードできるように開発を継続中（〈共同研究用IATサイト〉利用についてご関心をお持ちの研究者・大学院学生の方は，筆者まで直接ご連絡下さい）。

文献

Banaji, M. R., 2001, Implicit attitudes can be measured, In H. L. Roediger, III, J. S. Nairne, I. Neath, & A. Surprenant (Eds.), *The nature of remembering: Essays in honor of Robert G. Crowder*, Washington, DC: American Psychological Association, 117-150.

Bargh, J.A., 1997, The automaticity of everyday life, In R.S.Wyer, Jr. (Ed.), *Advances in social cognition* (Volume X.), Mahwah, NJ: Lawrence Erlbaum Associates, 1-61.

Blair, I.V., J.E. Ma & A.P. Lenton, 2001, Imagining stereotypes away: The moderation of implicit stereotypes through mental imagery, *Journal of Personality and Social Psychology*, 81: 828-841.

Collins, A.M. & E.F. Loftus, 1975, A spreading activation theory of semantic processing, *Psychological Review*, 82: 402-428.

Dasgupta, N., A.G. Greenwald, 2001, On the malleability of automatic attitudes: Combating automatic prejudice with images of admired and disliked individuals, *Journal of Personality & Social Psychology*, 81: 800-814.

Devine, P. G., 1989, Stereotyping and prejudice: Their automatic and controlled components, *Journal of Personality and Social Psychology*, 56: 5-18.

Greenwald, A. G., D. E. McGhee & J. K. L. Schwartz, 1998, Measuring individual differences in implicit cognition: The implicit association test, *Journal of Personality and Social Psychology*, 74: 1464-1480.

Greenwald, A. G. & B. A. Nosek, 2001, Health of the Implicit Association Test at Age 3, *Zeitschrift für Experimentelle Psychologie*, 48: 85-93.

Greenwald, A. G., B. A. Nosek & M. R. Banaji, 2003, Understanding and Using the Implicit Association Test: I. An Improved Scoring Algorithm, *Journal of Personality and Social Psychology*, 85: 197-216.

Lane, K. A., J. P. Mitchell & M.R. Banaji, 2001, Formation of implicit attitudes: Direct experience not required, Presented at the 2001 conference for the Society of Personality and Social Psychology, San Antonio, TX.

潮村公弘, 2016,『自分の中の隠された心――非意識的態度の社会心理学』サイエンス社(セレクション社会心理学 -29)。

潮村公弘, 2015,「潜在連合テスト (IAT) の実施手続きとガイドライン――紙筆版 IATを用いた実習プログラム・マニュアル」『対人社会心理学研究』15:31-38.

Shiomura, K, 2002, The effect of the attitudes shown by the respondent in self-disclosure on implicit and explicit self-esteem of the discloser: Investigation of the sensitive measure for clinical counseling, Presented at the X X V International Congress of Applied Psychology, Singapore.

Thurstone, L.L., 1928, Attitudes can be measured, *American Journal of Sociology*, 33: 529-554.