RUNNING HEAD: The Category-Focus IAT

The Category-Focus Implicit Association Test

Frank Siebler[1], Roberto González[2], Gabriela Ordóñez[2], Gerd Bohner[3], Andrés Haye[2], David Sirlopú[2], Andres Millar[2], Pablo de Tezanos-Pinto[2], and David Torres[2]

[1]Universitetet i Tromsø

[2]Pontificia Universidad Católica de Chile

[3]Universität Bielefeld

Abstract

We propose the Category-Focus IAT (CF-IAT) as an instrument to measure the implicit associations of single concepts. The CF-IAT aims to direct respondents' attention to a subset of the experimental materials. In a first study using the CF-IAT, Chilean adolescents ($\underline{N}$ = 49), members of either the indigenous minority (Mapuche) or the non-indigenous majority, completed CF-IATs assessing implicit attitudes toward these groups. Results revealed, in both groups, neutral implicit evaluation of the ingroup, but negative implicit evaluation of the outgroup. Process evidence suggests that the CF-IAT's manipulation of attentional focus was successful.

The Category-Focus Implicit Association Test

Research on social cognition distinguishes between <u>explicit cognition</u>, which is accessible to introspection, and <u>implicit cognition</u>, which reflects introspectively unidentified or inappropriately identified past experience (Greenwald & Banaji, 1995). A popular measure of implicit cognition is the Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998). It assesses the <u>difference</u> in association strength between each of two target concepts (e.g., "Self" and "Other") and an attribute dimension (e.g., positive or negative evaluation). Because of their relative (or comparative) nature, IAT scores are susceptible to apparently minor variations in the target concepts. For instance, implicit self-esteem scores, as assessed by a "Self versus Other" IAT, differed significantly if the "Other" concept referred to a close friend, rather than to an unspecified other (Karpinski, 2004; see also Brendl, Markman, & Messner, 2001). Several IAT-like <u>single-concept measures</u> are available that may mitigate the relativity issue. Below we review these instruments. Then we propose the Category-Focus IAT (CF-IAT), a new single-concept measure that combines the strengths of currently available instruments.[1] Finally, we report CF-IAT results from a study on implicit intergroup attitudes.

<u>Measures of implicit association</u>

Two aspects are most relevant for the present purpose, namely the numbers of stimulus categories that are (a) used in the stimulus set, and (b) addressed in the task instructions.

<u>The Standard IAT.</u> The IAT (Greenwald et al., 1998) measures the differential association of two target concepts (e.g. "Blacks" versus "Whites") with positive and negative evaluation (e.g. "pleasantness" versus "unpleasantness"). Stimuli representing these four categories are shown on the computer screen, one at a time. Participants respond by pressing one of two keys, according to task instructions that vary between trial blocks. In two critical blocks, task instructions assign two categories (both a target concept and an evaluation) to each response key. For instance, in the first block, participants may be asked to press the left-hand key if <u>a White face or a pleasant word</u> appears, but to press the right-hand key if <u>a Black face or an unpleasant word</u> appears. In the second

block, the assignment of target concepts (but not of evaluations) to response keys is reversed. If people associate one kind of faces more easily with pleasantness than with unpleasantness, then they are likely to complete one of the blocks more quickly than the other. The response speed difference between blocks defines a respondent's IAT score.

A strength of the IAT stems from the use of stimuli from an even number of categories (four: two target concepts, and two evaluations). This design allows for a well-balanced composition of the set of stimuli presented to participants. Both within and across blocks, three dimensions of the stimulus set can simultaneously be held constant: (a) the number of stimuli per category, (b) the presentation frequency of each stimulus, and (c) the proportions of left-hand and right-hand responses required. Further, because each block uses the same stimulus set, taking a difference score between blocks automatically removes undesired sources of variance (e.g., effects of variability in the word length of stimuli). Thereby, IAT scores are protected against a range of threats to validity related to the composition of the stimulus set.

A limitation of the IAT arises from the use of all four categories in the task instructions. By referring to all categories, the implicit evaluation of two target concepts is measured concurrently, yielding a single score that cannot be further decomposed (for discussion see Karpinski, 2004; Nosek & Banaji, 2001). Concurrent assessment is unique to the Standard IAT (see Table 1).

Single Concept IATs. Single-concept IATs use three categories in both the stimulus set and the task instructions. The Single-Target IAT (ST-IAT; Wigboldus, Holland, & Van Knippenberg, 2006) has a target concept share a response key with positive evaluation in one of the blocks (e.g., Islamic & positive versus negative), but with negative evaluation in the other (e.g., positive versus Islamic & negative). The Single-Attribute IAT (Penke, Eichstaedt, & Asendorpf, 2006) and the Single-Category IAT (Karpinski & Steinman, 2006; see also Bohner, Siebler, González, Haye, & Schmidt, 2008) also use stimuli from three categories in each block, address all categories in the task instructions, and change the reponse key assignment for one of the categories between blocks.

While solving the issue of concurrent assessment, the omission of a fourth category in the single-concept IATs comes at a cost: With an uneven number of categories, the stimulus set cannot be balanced on the same number of dimensions as was possible with the Standard IAT. First, if the proportions of left-hand versus right-hand responses shall be held constant in a block, the category not sharing a key must be represented by a greater number of stimuli. Second, the category requiring greater presentation frequency of its stimuli is a different one in each block. Consequently, the composition of the stimulus set differs between blocks. Taking a difference score between blocks therefore removes undesired variance to a lesser degree than in the Standard IAT. [2]

The Go/No-go Association Task. In the Go/No-go Association Task (GNAT; Nosek & Banaji, 2001), participants respond by either pressing ("go") or not pressing ("no-go") a single response key. To determine "no-go" decisions, response deadlines are applied, after which the experiment proceeds automatically. Importantly, the GNAT decouples the number of categories used in the stimulus set from the number of categories addressed in the task instructions: Task instructions name only those categories that require the "go" response. For example, in one block, participants were asked to press the key if a stimulus belonged to one of two categories (fruit & good), but not otherwise. By the actual composition of the stimulus set, the "no-go" response was thus tacitly assigned to two stimulus categories (bugs & bad) that remained unaddressed. In the next block, the assignment of evaluations to response options was reversed (fruit & bad versus, again tacitly, bugs & good). The performance difference between blocks was used as an index of the implicit attitude toward the directly addressed target concept (fruit).

Thus, while measuring implicit attitudes towards single objects, the GNAT maintains the Standard IAT's balanced stimulus set. Key to the GNAT procedure is a division of categories into foreground versus background categories, by either addressing or not addressing a given category in the task instructions. Rather than removing stimulus categories from the experimental materials, the GNAT removes categories from respondents' focus of attention.

The Category-Focus IAT

Our review identified the GNAT as the single-concept measure of greatest methodological rigor. However, different from the IAT paradigm, the GNAT paradigm uses response deadlines. These are not trivial to determine (see Nosek & Banaji, 2001). Moreover, in the GNAT paradigm, response latency data are systematically unavailable for "no-go" trials. A complete set of response latencies, as provided by the IAT paradigm, appears desirable because it enables additional insight into cognitive processes. We shall demonstrate this in our Results section.

From these considerations, we designed the CF-IAT to combine the strengths of the GNAT paradigm and the IAT paradigm. The CF-IAT uses balanced stimulus sets (comprising stimuli from four categories) and requires an active response to each stimulus. Its task instructions address a subset of the categories in the stimulus set. Specifically, whereas the categories sharing one of the response keys are addressed by their names (e.g., Black & pleasant), the categories sharing the other key are addressed by a summary label (e.g., Other). An implicit attitude score is computed as the performance difference between two blocks (e.g., Block 1: Black & pleasant vs. Other; Block 2: Black & unpleasant vs. Other).

The present study

We applied the CF-IAT in a larger study, conducted in 2002 and 2003, on the intergroup attitudes of non-indigenous Chileans and an indigenous Chilean minority, the Mapuche (Haye et al., in press). The relation between these groups is characterized by mutual prejudice and a history of conflict (Saiz, 1986, 2002). Therefore, members of either ethnic group should implicitly prefer their respective ingroup over the outgroup (see Greenwald et al., 1998, Study 2). Our participants completed two CF-IATs each. The same stimulus sets were used in both, but task instructions focused participants' attention on either the ingroup or the outgroup. We expected the mere variation of category focus to reveal a pattern of implicit attitudes such that ingroup CF-IAT scores should be more favorable than outgroup CF-IAT scores. This prediction was tested against a null-hypothesis whereby the two CF-IAT scores might not differ from each other, or might show a different pattern of means than predicted.

Method

Participants

Fifty eight students from four schools in Santiago, Chile, participated in the study. They attended courses for students of age 12-14 years. The data of nine participants were discarded due to an excessive number of overly quick responses (see below). The final sample comprised 49 participants (27 indigenous, 22 non-indigenous).

Design and variables

Participant origin (ethnicity: indigenous vs. non-indigenous) was treated as a between-subjects factor. A separate CF-IAT was conducted toward each target group, constituting the two levels of a within-subjects factor (category focus: on the Indigenous vs. on the Non-Indigenous). Note that the latter factor reflects the CF-IAT's core mechanism, the variation of task instructions. Within each CF-IAT, critical blocks were run twice. That served to familiarize participants with the task, but also allowed to compute two CF-IAT scores towards the same target; subsequent analyses treated the dual scores as the two levels of a within-subjects factor (block type: practice vs. test). The dependent variables were participants' four (2 target groups X 2 block types) CF-IAT scores each.

Materials and procedure

Stimulus materials. The stimuli were drawn from a pretested pool of pleasant words, unpleasant words, indigenous stimuli, and non-indigenous stimuli (six to seven items each). Example words are, pleasant: peace, love; unpleasant: war, sadness. The stimuli referring to social groups were typical surnames (e.g., indigenous: Huilcaleo; non-indigenous: Morales) and digital photographs of indigenous and non-indigenous Chileans.

Trial blocks. Participants completed ten trial blocks (see Table 2). A block's assignment of categories to response keys was indicated by initial instructions, and by labels in the top-left and top-right corners of the screen. Importantly, whereas we actually presented stimuli from all four categories in each critical block, instructions and on-screen labels directly addressed only a subset

(two of four) of these categories. Stimuli were presented one by one, centered on the computer screen. Incorrect responses were indicated by the letter "X" in red color and needed to be corrected. Correct responses triggered a 250ms blank-screen interval, followed by the next stimulus. Practice blocks comprised 28 stimuli; test blocks comprised 40 stimuli.

Scoring. CF-IAT scores depict the implicit evaluation of the target group that was addressed in the task instructions. We computed separate scores from practice block data and test block data, respectively. Four scores were computed: Non-Indigenous target / practice (Block 3 minus Block 7), Non-Indigenous target / test (Block 4 minus Block 8), Mapuche target / practice (Block 9 minus Block 5), and Mapuche target / test (Block 10 minus Block 6).

The computation of scores followed the improved IAT scoring algorithm (Greenwald, Nosek, & Banaji, 2003), yielding a measure called D. In line with recommendations, we removed cases with more than ten percent of overly quick responses (i.e., faster than 300 ms). On the resulting CF-IAT D scores, a positive sign indicates positive evaluation of the target group, whereas a negative sign indicates negative evaluation of the target group.

## Results

### Category-Focus IAT

Reliability of measures. Practice block and test block CF-IAT scores correlated positively and significantly, both when the target group was the ingroup, $r(47) = .30$, $p = .037$, and when the target group was the outgroup, $r(47) = .28$, $p = .049$. For comparison, the correlation of scores towards different target groups was small and nonsignificant (practice blocks: $r = .10$; test blocks: $r = .05$), suggesting that the two CF-IAT scores did assess distinct attitudes as intended.

Difference between conditions. We entered participants' four CF-IAT scores into a mixed-model ANOVA with participant ethnicity (indigenous vs. non-indigenous) as between-subjects factor, and repeated measures on block type (practice vs. test) as well as category focus (on the Mapuche vs. on the Non-Indigenous). See Table 3 for means and standard deviations. The analysis

revealed a significant two-way interaction of participant ethnicity and category focus, $F(1,47) = 17.31$, $p < .001$, $\underline{MSE} = 0.85$; all other $\underline{F} < 1.5$.

To follow up on the interaction we repeated the analysis within participant groups. For <u>non-indigenous participants</u>, scores were comparatively more positive when category focus was on the Non-Indigenous rather than the Mapuche, $\underline{F}(1,21) = 13.07$, $\underline{p} = .002$, $\underline{MSE} = 0.68$; all other $\underline{F} < 1$. For <u>indigenous participants</u>, scores were comparatively more positive when the target group were the Mapuche rather than the Non-Indigenous, $\underline{F}(1,26) = 5.98$, $\underline{p} = .022$, $\underline{MSE} = 1.00$; all other $\underline{F} < 1$. Thus, the predicted pattern of intergroup attitudes was observed in both subsamples.

<u>Difference from zero.</u> CF-IAT scores were submitted to $\underline{t}$-tests for difference from zero (neutral). For either participant group, scores towards the respective ingroup were neutral, whereas scores towards the respective outgroup were unequivocally negative (see Table 3).

<u>Process evidence</u>

A final set of analyses addressed whether participants focused on the intended subset of categories. Directly naming some categories in the task instructions, while referring to others only indirectly (with the summary label "Other"), should render the former categories more accessible in memory than the latter. Heightened accessibility of a category, in turn, should facilitate responding to its associated stimuli. To test for the expected response facilitation, we computed two response latency means. The first mean averaged across those trials where the stimulus' category was named directly. The second mean averaged across the remaining trials (where the stimulus needed to be classified as "Other"). A difference score was then derived by subtracting the former mean score from the latter, and was transformed into a $\underline{D}$ score. On the resulting score, a positive (negative) sign indicates that participants responded more (less) quickly if stimuli belonged to a category that was addressed directly, rather than by the summary label. Separate scores were computed for each of the four (block type x target group) within-subjects conditions.

<u>Difference between conditions.</u> We submitted participants' four scores each to a mixed-model ANOVA with participant ethnicity (indigenous vs. non-indigenous) as between-subjects

factor, and repeated measures on block type (practice vs. test) as well as category focus (on the

Mapuche vs. on the Non-Indigenous). As the only effect approaching significance in this analysis,

test blocks tended to have more positive scores than practice blocks, $\underline{F}(1,47) = 3.54$, $\underline{p} = .066$,

$\underline{MSE} = 0.06$; all other $\underline{F} < 1.9$.

      <u>Difference from zero.</u> After collapsing across factors that had not shown an effect in the

ANOVA, scores were submitted to $\underline{t}$-tests for difference from zero. Whereas practice block scores

($\underline{M} = 0.02$, $\underline{SD} = 0.19$) did not differ significantly from zero, $\underline{t} < 1$, <u>ns.</u>, test block scores ($\underline{M} = 0.08$,

$\underline{SD} = 0.16$) did, $\underline{t}(48) = 3.74$, $\underline{p} < .001$. Thus, for test blocks, the category focus manipulation

demonstrably resulted in an expected response facilitation.

<div align="center">Discussion</div>

      The CF-IAT is a new measure of single concepts' implicit associations; it was designed to

integrate the strengths of existing instruments. A first study with the CF-IAT revealed the predicted

pattern of implicit intergroup attitudes in a sample of Chilean students. The CF-IAT aims to direct

respondents' attention to a subset of the categories that comprise the stimulus set. The process

evidence suggests that respondents did in fact focus their attention as prescribed by task instructions,

and did so more strongly with greater practice. CF-IAT scores computed from practice blocks and

test blocks showed a significant but low correlation. This is likely to underestimate the instrument's

internal consistency, which may increase with more sophisticated analyses, or (as suggested by the

process evidence) when giving respondents more practice. Finally, we should point out that the CF-

IAT does not assess implicit associations in an "absolute" or "context-free" way. For the time being,

however, the CF-IAT appears a promising tool for the assessment of single-concept associations.

Endnotes

[1] Sriram and Greenwald (2009) independently developed a similar instrument, the Brief Implicit Association Test.

[2] It should be noted that the authors of single-concept IATs devised measures to reduce the impact of confounds, for example in the calculation of the score.

References

Bohner, G., Siebler, F., González, R., Haye, A., & Schmidt, E. A. (2008). Situational flexibility of in-group-related attitudes: A single category IAT study of people with dual national identity. Group Processes and Intergroup Relations, 11, 301-317.

Brendl, C. M., Markman, A. B., & Messner, C. (2001). How do indirect measures of evaluation work? Evaluating the inference of prejudice in the Implicit Association Test. Journal of Personality and Social Psychology, 81, 760–773.

Greenwald, A. G., & Banaji, M. R. (1995). Implicit social cognition: Attitudes, self-esteem, and stereotypes. Psychological Review, 102, 4-27.

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.

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.

Haye, A., González, R., Ordóñez, G., Bohner, G., Siebler, F., Sirlopú, D., Millar, D., De Tezanos-Pinto, P., & Torres, D. (in press). System-perpetuating asymmetries between explicit and implicit intergroup attitudes among indigenous and non-indigenous Chileans. Asian Journal of Social Psychology.

Karpinski, A. (2004). Measuring self-esteem using the Implicit Association Test: The role of the other. Personality and Social Psychology Bulletin, 30, 22-34.

Karpinski, A. & Steinman, R. B. (2006). The single category implicit association test as a measure of implicit social cognition. Journal of Personality and Social Psychology, 91, 16-32.

Nosek, B. A., & Banaji, M. R. (2001). The go/no-go association task. Social Cognition, 19, 625-664.

Penke, L., Eichstaedt, J., & Asendorpf, J. B. (2006). Single-attribute implicit association tests (SA-IAT) for the assessment of unipolar constructs: The case of sociosexuality. <u>Experimental Psychology, 53</u>, 283-291.

Saiz, J. L. (1986). Estereotipos adscritos al indígena mapuche por adultos no mapuches de Chile meridional. <u>Revista Interamericana de Psicología, 20</u>, 55-68.

Saiz, J. L. (2002). Atribución de estereotipos: Los indígenas mapuches que perciben los chilenos. In J. F. Morales, D. Páez, A. L. Kornblit, & D. Asún (Eds.), <u>Psicología social</u> (pp. 145-151). Buenos Aires: Pearson Education.

Sriram, N., & Greenwald, A. G. (2009). The Brief Implicit Association Test. <u>Experimental Psychology, 56</u>, 283-294.

Wigboldus, D. H. J., Holland, R. W., & Van Knippenberg, A. (2006). <u>Single target implicit associations.</u> Manuscript under review.

Table 1

Properties of Measures of Implicit Association

| Measure | No. of distinct categories | | No. of response keys | Assessment of 2nd target |
|---|---|---|---|---|
| | Stimulus set | Task instructions | | |
| Standard IAT | 4 | 4 | 2 | concurrently |
| Single Concept IATs | 3 | 3 | 2 | sequentially |
| Go/No-go AT | 4 | 2 | 1 | sequentially |
| Category-Focus IAT | 4 | 2 | 2 | sequentially |

Notes. The table depicts properties of measures of implicit association. Column "Stimulus set" shows the number of distinct categories that comprise the stimulus set. Column "Task instructions" shows the count of categories in the stimulus set that are directly addressed in task instructions. Column "No. of response keys" shows the count of response keys available to participants. Column "Assessment of 2nd target" reflects whether attitudes toward first and second target object (if any) are assessed simultaneously or one after the other. AT = association task; IAT = implicit association test.

Table 2

Trial-Block Order

| Block No. | Left-hand label | Right-hand label |
|:---:|:---:|:---:|
| 1 | MAPUCHE | NON-INDIGENOUS |
| 2 | PLEASANT | UNPLEASANT |
| 3 | NON-INDIGENOUS or UNPLEASANT | OTHER |
| 4 | NON-INDIGENOUS or UNPLEASANT | OTHER |
| 5 | MAPUCHE or PLEASANT | OTHER |
| 6 | MAPUCHE or PLEASANT | OTHER |
| 7 | NON-INDIGENOUS or PLEASANT | OTHER |
| 8 | NON-INDIGENOUS or PLEASANT | OTHER |
| 9 | MAPUCHE or UNPLEASANT | OTHER |
| 10 | MAPUCHE or UNPLEASANT | OTHER |

Notes. Ten trial blocks were completed. Eight of these were critical: four practice blocks (3, 5, 7, and 9) as well as four test blocks (4, 6, 8, and 10).

Table 3

Category-Focus IAT Scores as a Function of Participant Ethnicity and Target Group

|  | Participant Ethnicity | |
|---|---|---|
|  | Mapuche | Non-Indigenous |
| **Target Group** | | |
| **Practice** | | |
| Mapuche | -.02  (.87) | -.68* (.90) |
| Non-Indigenous | -.36+ (.96) | .01  (.97) |
| **Test** | | |
| Mapuche | -.03  (.77) | -.50* (.68) |
| Non-Indigenous | -.62** (.68) | .08  (.68) |
| **Practice & Test** | | |
| Mapuche | -.02  (.63) | -.59** (.61) |
| Non-Indigenous | -.49* (.69) | .05  (.70) |

Notes. Cell entries depict D-scores on the Category-Focus-IAT measures, separately for practice blocks (top panel), test blocks (middle panel), and practice blocks plus test blocks combined (bottom panel). Positive (negative) scores indicate positive (negative) implicit evaluation of the target group. Figures in parentheses are standard deviations. Difference from zero: + $p < .10$, * $p < .01$, ** $p < .001$. Cell N, from left to right: 27, 22.