

The Psychological Record, 2009, 59, 621–640

THE IMPLICIT RELATIONAL ASSESSMENT PROCEDURE (IRAP) AS A MEASURE OF IMPLICIT RELATIVE PREFERENCES: A FIRST STUDY

Patricia Power, Dermot Barnes-Holmes, Yvonne Barnes-Holmes

National University of Ireland, Maynooth

Ian Stewart

National University of Ireland, Galway

The Implicit Relational Assessment Procedure (IRAP) was designed to examine implicit beliefs or attitudes. In Experiment 1, response latencies obtained from Irish participants on the IRAP showed a strong preference for Irish over Scottish and American over African. In contrast, responses to explicit Likert measures diverged from the IRAP performance in indicating Irish equally likeable to Scottish and African more likeable than American. Using a similar IRAP, Experiment 2 showed that participants from the United States showed strong implicit preferences for American over Irish, Irish over Scottish, and Scottish over African; the explicit Likert measures again diverged from the IRAP. The findings provide preliminary support for the IRAP as a useful measure of implicit beliefs.

Keywords: implicit relational assessment procedure, implicit beliefs, social attitudes, adult participants, patterns of preference

Traditional methods for assessing or measuring what people think or believe have typically involved self-report questionnaires, which, it has been argued, may be subject to a number of biases. For example, individuals may answer in a way they judge to be the most socially acceptable (e.g., for socially sensitive topics). Furthermore, individuals may not even be aware of their own beliefs, or if awareness does exist, they make efforts to conceal their views if they are deemed to be socially undesirable. As a result, increasing attention has been devoted to studying what are described as *implicit* beliefs or attitudes. Intense debate surrounds the nature of these implicit cognitions, but one of the most widely cited definitions was provided by Greenwald and Banaji (1995), who stated that implicit attitudes are “introspectively

Patricia Power, Dermot Barnes-Holmes, Yvonne Barnes-Holmes, Department of Psychology, National University of Ireland, Maynooth, Ireland. Ian Stewart, Department of Psychology, National University of Ireland, Galway, Ireland.

Preparation of the current article was supported by a postgraduate scholarship awarded to Patricia Power, from the Irish Research Council for Science, Engineering and Technology.

Correspondence concerning this article should be addressed to Patricia Power (email: patricia.m.power@gmail.com) or Dermot Barnes-Holmes (email: Dermot.Barnes-Holmes@nuim.ie).

unidentified (or inaccurately identified) traces of past experience that mediate favorable or unfavorable feeling, thought, or action toward social objects” (p. 8).

The study of implicit cognition has led to the development of a range of different experimental procedures that are designed to reveal implicit beliefs pertaining to a particular domain. These types of measures are believed to address some of the weaknesses inherent in traditional self-report or explicit measures. Presently, the *Implicit Association Test* (IAT; Greenwald, McGhee, & Schwartz, 1998) is the most popular measure of implicit beliefs or attitudes. The IAT is based on the assumption that it should be easier to map two concepts onto a single response when those concepts are associated in memory than when the concepts are unrelated. In a seminal study, Greenwald et al. presented participants with the names of flowers (e.g., “tulip”), the names of insects (e.g., “beetle”), pleasant words (e.g., *love*), and unpleasant words (e.g., *ugly*). It was assumed that the concept of flower and the attribute “pleasant” are likely associated in memory, as are the concept of insect and the attribute “unpleasant.” The researchers further reasoned that responses should be faster when response keys are assigned to associated stimuli than when they are not. The results of the experiment supported this prediction; response latencies were shorter when both “flower” and “pleasant” were assigned to one key and “insect” and “unpleasant” were assigned to a second key than when unassociated stimuli were assigned to the same keys (key 1 = flower + unpleasant; key 2 = insect + pleasant).

Greenwald and colleagues have now published a range of studies showing that the IAT appears to reveal implicit attitudes that participants typically deny, such as racial stereotypes, when explicit measures such as questionnaires and open-ended interviews are used (see Nosek, Greenwald, & Banaji, 2006, for a review). For example, IAT research has shown that many participants who claimed *not* to hold racist attitudes nevertheless responded more quickly and more accurately when asked to categorize names typical of White persons with positive words and names typical of Black persons with negative words, than when asked to categorize White with negative and Black with positive (see Greenwald et al., 2002). This basic IAT effect has now been replicated many times, and it has become an increasingly popular method for assessing implicit attitudes and dysfunctional beliefs (e.g., De Houwer, 2002).

In providing a structural analysis of the IAT, De Houwer (2003a) argued that the test involves manipulating relevant and irrelevant stimulus response (S-R) compatibility.¹ For present purposes we will focus only on the former manipulation because De Houwer also concluded that it is primarily responsible for the IAT effect. *Relevant S-R compatibility* refers to the physical or functional similarity between a task-relevant stimulus feature and the required response. During compatible trials in a traditional S-R compatibility task, a participant might be required to press a left key when a dot appears on the left and a right key when a dot appears on the right, but during incompatible trials the opposite performance is required: Left dot = press right and right dot = press left. In effect, the two types of trials differ in terms of the level of *relevant S-R compatibility*, and results indicate that compatible trials typically produce a superior performance.

1 De Houwer also provides structural analyses of the Emotional Stroop, the Extrinsic Affective Simon Task, and Evaluative Priming.

According to De Houwer (2003a), the IAT is functionally similar to a traditional S-R compatibility task because the level of relevant S-R compatibility varies across compatible and incompatible trials (i.e., compatible task = positive or positive vs. incompatible task = positive or negative). In an IAT in which flower names, insect names, positive words, and negative words are presented, the relevant stimulus feature is the category to which the presented stimulus belongs. The valence of the relevant feature of the stimulus is positive on some trials (flower and positive) and negative on others (insect and negative), and thus responses also receive an extrinsic valence because of the response assignments. On compatible trials (e.g., “Press the left key for positive words and flower names”), the response key will be linked with both the concept of positive and the concept of flower, and will therefore adopt a clearly positive extrinsic meaning. On incompatible trials, however, the response key will be linked with both the concept of positive and the concept of insect, and thus the extrinsic valence of the response will be unclear. The IAT thus involves switching between tasks that involve compatible and incompatible S-R relations.

Another implicit measure that has recently been offered also appears to involve manipulating relevant S-R compatibility. This measure is called the Implicit Relational Assessment Procedure (IRAP; Barnes-Holmes, Barnes-Holmes, Power, Hayden, Milne, & Stewart, 2006), which emerged from a modern behavior-analytic account of human language and cognition known as Relational Frame Theory (RFT; Hayes, Barnes-Holmes, & Roche, 2001). The core postulate of RFT is that the basic units of language and cognition are relational, and thus the procedures that have emerged from research in this area tend to focus on stimulus relations and relational networks (e.g., Barnes-Holmes, Healy, & Hayes, 2000). Consistent with this approach, the IRAP focuses on stimulus relations, and structurally it could be described as a task that manipulates relevant relational-response (R-R), rather than S-R, compatibility. Within the behavior-analytic tradition, however, relations are defined as stimuli (Catania, 198), and thus the distinction may be more procedural than process based.

The IRAP is best described as a combination of the IAT and an earlier RFT-based procedure known as the Relational Evaluation Procedure (REP), which has been used to examine relational responding across a range of studies (Cullinan, Barnes, & Smeets, 1998; Cullinan, Barnes-Holmes, & Smeets, 2000, 2001; O’Hora, Barnes-Holmes, Roche, & Smeets, 2004; O’Hora, Pelaez, & Barnes-Holmes, 2005; Stewart, Barnes-Holmes, & Roche, 2002, 2004). Like the REP, the IRAP involves presenting specific relational terms (e.g., *similar*, *opposite*, *better*, *worse*) so that the properties of the relations among the relevant stimuli can be assessed. And similar to the IAT, the IRAP involves asking participants to respond quickly and accurately in ways that are either consistent or inconsistent with their pre-experimentally established verbal relations. The basic hypothesis is that average response latencies should be shorter across blocks of consistent relative to inconsistent trials. Or, in other words, participants should respond more rapidly to relational tasks that reflect their current beliefs than to tasks that do not.

As stated above, like the IAT, the IRAP involves manipulating relevant S-R (or R-R) compatibility. That is, the IRAP requires that participants respond to specific stimulus relations in a manner that is either consistent

or inconsistent with established verbal relations. Thus, for example, a participant might be asked to respond *True* to the target statement “Mouse bigger than a fly” for a consistent trial, but respond *False* for an inconsistent trial. The basic assumption is that when such relational responses are made under time pressure, average latencies will be faster for consistent relative to inconsistent trials. This prediction is based on the structural similarity between the IAT and IRAP, in that both involve manipulating relevant S-R compatibility. We recognize that a prediction based on a structural analysis does not involve specifying the psychological processes involved in producing a behavioral effect, but it may help to isolate the critical independent variable or variables. In any case, in the General Discussion we shall consider the issue of psychological process, and how the RFT approach to implicit cognition differs from the mainstream perspective.

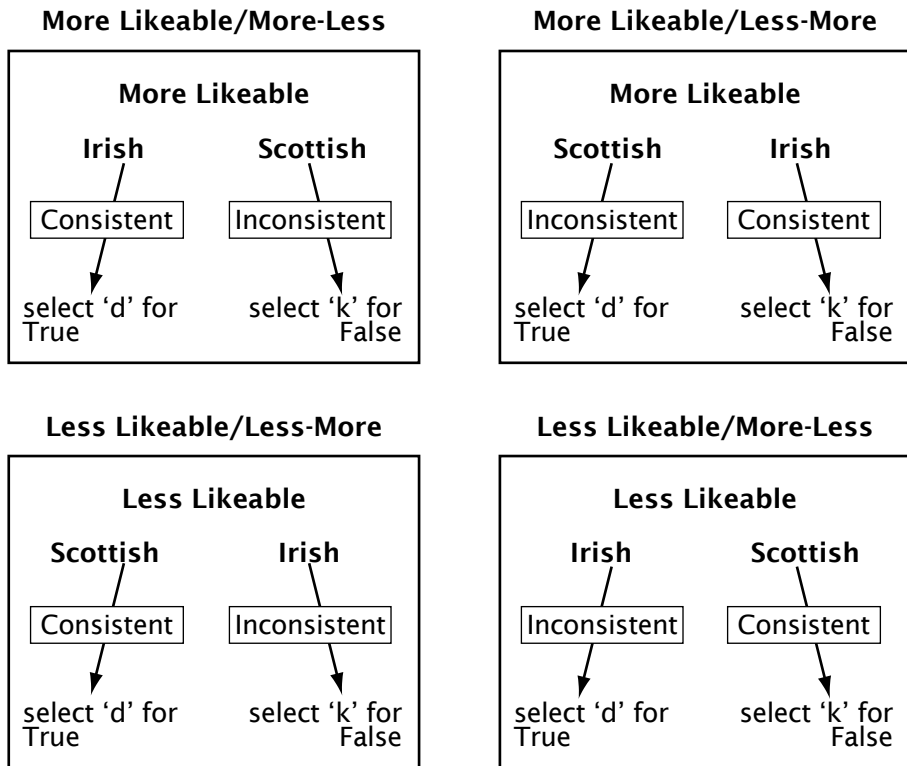


Figure 1. Examples of the four IRAP trial types employed in Experiment 1. The sample (More Likeable or Less Likeable), target word pairs (Irish-Scottish, Scottish-American, etc.), and response options (True and False) appeared simultaneously on each trial. Arrows with superimposed text boxes indicate which responses were deemed consistent or inconsistent (boxes and arrows did not appear on-screen). Selecting the consistent response option during a consistent block, or the inconsistent option during an inconsistent block, cleared the screen for 400 ms before the next trial was presented; if the inconsistent option was chosen during a consistent block, or the consistent option during an inconsistent block, a red X appeared on-screen until the participant gave the alternative response.

One of the aims of the current study was to determine if the IRAP could be used to identify a socially sensitive attitude that is readily predicted on the basis of known psychological principles but may well be hidden when assessed with an explicit measure. A second aim of the current study was to determine if the IRAP could be used to assess beliefs that involve comparative relations. Experiment 1 sought to test the comparative relations “More Likeable” and “Less Likeable” using socially valenced target stimuli. On each trial of the IRAP in Experiment 1, one of the two samples “More Likeable” and “Less Likeable” was presented with a target stimulus consisting of two words that referred to different national groups (e.g., *Irish-Scottish*, *Scottish-American*, *American-African*), and vice versa (e.g., *Scottish-Irish*). Pilot work had shown that Irish individuals, given a choice among Scottish, American, and African, typically rated Scottish as most similar to Irish, with American ranked next, followed by African. Research has shown that individuals generally show a preference for their own group over other groups (Tajfel, 1982), and that relative preferences for other groups are determined, in part, by the extent to which the other groups are perceived to be more or less similar to one’s own group (see Cota, Evans, Dion, Kilik, & Longman, 1995; Goethals & Darley, 1977). On the basis of the ranking obtained in our pilot work, we predicted that Irish participants in the current study might show a preference for Irish over Scottish, Scottish over American, and American over African.

To determine if this pattern of preferences would be obtained using an explicit measure, 16 Irish participants were asked to rate on four separate Likert scales the extent to which they found Irish, Scottish, American, and African people likeable or unlikeable. The scales ranged from -6, indicating *extremely unlikeable*, to +6, indicating *extremely likeable*. The IRAP was then employed to determine if the participants’ implicit relative preferences for the four groups would differ from their explicit responses to the Likert scales. Trials were deemed consistent or inconsistent in accordance with the prediction that Irish participants would show a greater liking for Irish over Scottish, Scottish over American, and American over African. For example, response latencies should be shorter when Irish participants have to respond “True” to “Irish more likeable than Scottish” and “Scottish less likeable than Irish” than when they have to respond “False” on these trials. Figure 1 presents examples of the four IRAP trial types employed in Experiment 1.

Experiment 1

Method

Participants

Sixteen Irish citizens (9 male and 7 female) agreed to participate in Experiment 1. Ages ranged from 18 to 35 years ($M = 28$ years). No financial payment or other inducements were offered for participation in the study. Eight participants each were assigned randomly to one of two conditions: consistent relations first or inconsistent relations first (the two conditions are described subsequently).

Setting

All of the experimental trials were conducted in a quiet room in the Department of Psychology at the National University of Ireland, Maynooth, where participants were seated comfortably and alone at a table in front of a personal computer. All participants completed the study on an individual basis. Following the instructional and practice phases of the experiment, the experimenter left the room and remained seated outside the door.

Apparatus and Materials

The IRAP was presented to participants on a Pentium 4 personal computer. Each IRAP trial presented one of two sample stimuli: "More Likeable" or "Less Likeable." On each trial, participants were also presented with one of six target word pairs (see Table 1). Finally, each trial presented two response options: "True" and "False."

Table 1
IRAP Stimuli in Experiment 1

Sample 1 More Likeable	Sample 2 Less Likeable
Response option 1 True	Response option 2 False
Targets deemed consistent with Sample 1 Irish Scottish Scottish American American African	Targets deemed consistent with Sample 2 Scottish Irish American Scottish African American

Four Likert scales were designed to assess the participants' explicit ratings of how likeable they found the four target nationalities (Irish, Scottish, American, and African). Each rating involved a 13-point scale that ranged from -6 (*extremely unlikeable*) to +6 (*extremely likeable*). For example, participants were asked to "Please indicate the extent to which you find American people to be likeable or unlikeable by circling the appropriate number."

Procedure

Phase 1: The explicit measure. Each participant was first presented with four Likert scales, which asked them to rate the likeability of Irish, Scottish, American, and African people, respectively. Following completion of the scales, participants were presented with a sheet explaining that the next part of the study would involve a computer-based task requiring them to make speedy and accurate responses that might or might not correspond to their own beliefs.

Phase 2: The IRAP. Before participants were exposed to the IRAP tasks, the program presented a series of instructions that described the tasks. On each IRAP trial, six words appeared simultaneously on the screen. A sample stimulus—either "More Likeable" or "Less Likeable"—appeared at the top, with a target word pair presented in the center and two response options, "True" and "False," at the bottom left- and right-hand corners of the screen (see Figure 1). All of the stimuli remained visible until the participant pressed one of the response keys. The task involved choosing one of the two response

options (“True” or “False”). To choose the term on the left, participants pressed the “d” key, and to choose the term on the right, participants pressed the “k” key (using their left and right index fingers, respectively). Choosing the response option that was deemed correct for that block of trials removed all six stimuli from the screen for 400 ms before the next trial was presented. Choosing the response option that was deemed incorrect for that block of trials produced a red X in the middle of the screen (immediately below the target word pair). The participant could not proceed to the next trial until he or she chose the correct response option (for that block of trials). In effect, all key presses other than the correct one were nonfunctioning.

All participants were exposed to eight blocks of 24 trials: two practice blocks followed by six test blocks. Within each block, the six target word-pairs (*Irish-Scottish*, *Scottish-American*, and *American-African*, and their reversed counterparts; see Table 1) were presented in a random order, with the constraint that each word pair was presented four times, twice in the presence of “More Likeable” and twice in the presence of “Less Likeable.” As described in the penultimate paragraph of the introduction, trials were deemed consistent when they were compatible with the predicted relative likeability of the target nationalities. That is, Irish participants should show a preference for Irish over Scottish, Scottish over American, and American over African. The reverse pattern applied to the inconsistent trials. The left-right position of the response options (“True” and “False”) alternated randomly across trials.

Consistent relations first. The following message was presented on-screen before the participants were exposed to each practice block: “This is a practice—errors are expected.” The first block of trials in the consistent-relations-first condition reinforced responses that were deemed consistent with the prediction that Irish participants might show a preference for Irish over Scottish, Scottish over American, and American over African. For example, given the sample “Less Likeable” and the target word pair *Scottish-Irish*, choosing the response option “True” immediately progressed the computer program to the next trial (after 400 ms). If, however, “False” was chosen, a red X appeared below the stimulus and the participant had to make the correct response by choosing “True” in order to progress to the next trial.

The second block of trials reinforced responses that were deemed inconsistent with the predicted relative preferences. For example, given the sample “More Likeable” and the target word pair *Scottish-Irish*, choosing the response option “True” immediately progressed the computer program to the next trial (after 400 ms). If, however, “False” was chosen, the red X appeared below the stimulus and the participant had to correct the error (choose “True”) to proceed. The Experimenter sat beside the participant during the first two practice blocks to answer any questions and to ensure that the task was fully understood. The experimenter then withdrew to another room for the remaining six blocks, which were presented in exactly the same format as the first two practice blocks. These test blocks alternated between consistent and inconsistent conditions. At no point was the participant informed which block was considered consistent or inconsistent.

At the end of each block, a feedback screen was presented that provided the percentage of correct responses and median response latency for that block of trials. Before each new block of trials, a message was presented

on-screen informing participants that in the next phase, “the previously correct and wrong answers are reversed.” Immediately before each of the six test blocks, the following text appeared on the screen: “This is a test—go fast; making a few errors is OK.” After completion of all eight blocks, the participant was thanked and appropriately debriefed.

Inconsistent relations first. Participants in the inconsistent-relations-first condition were exposed to the procedures described previously, except that the practice and test sequences alternated from inconsistent to consistent blocks.

Results and Discussion

Explicit Likert Analyses

A mean Likert rating for each of the four social groups was calculated across the 16 participants (Irish, $M = 3.38$, $SE = .364$; Scottish, $M = 3.44$, $SE = .302$; American, $M = 0.13$, $SE = .598$; African, $M = 1.81$, $SE = .41$). Irish and Scottish were rated as almost equally likeable, with African less likeable than the former two but considerably more likeable than American. The four ratings from each participant were entered into a one-way repeated measures ANOVA, which yielded a significant result, $F(15, 3) = 28.69$, $p < .0001$, $\eta_p^2 = .7$. Post hoc Fisher’s PLSD tests indicated that the ratings for each social group were significantly different from each other (all $ps < .0006$), with the exception of the Irish-Scottish comparison ($p = .88$). This pattern of preferences was not consistent with the prediction made in terms of perceived group similarity, in that (a) Irish participants did not rate Irish as more likeable than Scottish, and (b) African was rated as significantly more likeable than American. Would the IRAP show a similar pattern of preferences?

IRAP Preliminary Analyses

The primary datum was response latency defined as the time in milliseconds that elapsed between the onset of the trial and a correct response made by the participant. For the purposes of the current study, responses of more than 3,000 ms were recorded as 3,000 ms, and responses of less than 300 ms were recorded as 300 ms (see Greenwald, Nosek, & Banaji, 2003, for a detailed discussion of this and other algorithms that may be used for transforming response latency data generated by implicit measures).

To determine whether the order of testing (consistent-first vs. inconsistent-first) or the test exposure (Test 1, Test 2, or Test 3) had a significant impact on the IRAP performance, we calculated the overall adjusted mean latencies for each of the six test blocks for each participant; the overall means and standard errors are presented in Table 2. Within each pair of test blocks, the mean response latencies were consistently shorter for the consistent relative to the inconsistent trials. For example, the overall mean latency for consistent-first participants was 2,279 ms for the first test block (consistent) and 2,516 ms for the second (inconsistent); for the inconsistent-first condition the means were 2,407 ms for the first block (inconsistent) and 2,224 ms for the second (consistent).

The mean latencies were subjected to a $2 \times 2 \times 3$ mixed repeated measures ANOVA with test type (consistent vs. inconsistent) and test

exposure (1, 2, or 3) as repeated measures and test sequence (consistent-first vs. inconsistent-first) as a between-participants variable. Analyses indicated that the difference in latencies between consistent and inconsistent trials reached statistical significance, $F(1, 14) = 52.519, p < .0001, \eta_p^2 = .9$, but all other main and interaction effects did not (all $ps > .15$). Test-order and test-exposure variables were thus removed from all subsequent analyses.

Table 2
Overall Adjusted Mean Latencies and Standard Errors in Each Condition for Each Test Block

Test order	IRAP test type per block											
	Consistent						Inconsistent					
	Test 1		Test 2		Test 3		Test 1		Test 2		Test 3	
	M	SE	M	SE	M	SE	M	SE	M	SE	M	SE
Consistent first	2,279	100	2,295	84	2,221	115	2,516	105	2,537	101	2,492	132
Inconsistent first	2,224	133	2,127	154	2,155	148	2,407	115	2,532	119	2,406	131

Individual Pairs Analyses

The mean response latencies for each participant for consistent and inconsistent trials were calculated for each of the three target pairs (see Table 3). For each pair, response latencies were shorter for consistent relative to inconsistent trials. A 2 × 3 ANOVA was conducted with test type (consistent vs. inconsistent) and target pair (1-3) as repeated measures. Ignoring the previously reported main effect for test type, the effect for target pair was significant, $F(2, 28) = 47.533, p < .0001, \eta_p^2 = .8$, as was the interaction, $F(2, 28) = 5.746, p = .008, \eta_p^2 = .3$. Three separate ANOVAs were used to analyze consistent versus inconsistent latency differences for each pair. In each case, the difference proved to be significant (all $ps < .0001$, all $\eta_p^2 > .7$). Overall, therefore, each target pair produced the predicted IRAP effect.

Table 3.
Overall Mean Latencies and Standard Errors for Consistent and Inconsistent Trials for Each Target Pair in Experiment 1

Target pair	IRAP test type	M	SE
Irish-Scottish	Consistent	2,024	85
	Inconsistent	2,387	83
Scottish-American	Consistent	2,297	88
	Inconsistent	2,540	74
American-African	Consistent	2,328	85
	Inconsistent	2,567	79

Explicit-Implicit Correlations

To conduct correlational analyses of the IRAP and self-report data, we

calculated difference scores for each of the three pairs for each measure. For the IRAP data, this involved subtracting each participant's mean latency score for consistent trials from that for inconsistent trials for each pair. For the Likert data, each participant's rating for Scottish was subtracted from Irish, and similarly for American from Scottish, and African from American. Two of the three Spearman rank correlations were nonsignificant, although the effect for the American–African pair was negative (Irish–Scottish, $r_s = .208$, $p = .64$; American–African, $r_s = -.35$, $p = .13$). In contrast, the Scottish–American pair yielded a strong positive correlation between the IRAP and Likert measures ($r_s = .7$, $p = .006$).

Trial-Type Analyses

To explore any trial-type effects, the response latency data for each participant were calculated for each trial type for consistent and inconsistent trials (see Table 4). The data were subjected to a $2 \times 2 \times 2$ ANOVA, with sample, target type, and consistent versus inconsistent as repeated measures. The primary purpose of this analysis was to determine if IRAP test type (consistent vs. inconsistent) interacted with either of the other two variables (sample and target type), but it did not (all $ps > .12$). In fact, apart from the significant main effect for test type (recorded by the previous ANOVAs), the only significant result was a main effect for sample, $F(1, 14) = 34.845$, $p < .0001$, $\eta_p^2 = .7$. This result indicates that participants responded more quickly on trials that presented “More Likeable” than trials that presented “Less Likeable” ($M = 2,315$ ms, $SE = 44$, and $M = 2,484$ ms, $SE = 43$, respectively).

Table 4.
Overall Mean Latencies and Standard Errors for Each Trial Type for Consistent and Inconsistent Trials

Sample	Target	IRAP test type	<i>M</i>	<i>SE</i>
Less Likeable	Less Likeable to More Likeable	Consistent	2,341	86
Less Likeable	Less Likeable to More Likeable	Inconsistent	2,611	80
Less Likeable	More Likeable to Less Likeable	Consistent	2,418	88
Less Likeable	More Likeable to Less Likeable	Inconsistent	2,567	78
More Likeable	Less Likeable to More Likeable	Consistent	2,193	85
More Likeable	Less Likeable to More Likeable	Inconsistent	2,453	83
More Likeable	More Likeable to Less Likeable	Consistent	2,136	73
More Likeable	More Likeable to Less Likeable	Inconsistent	2,479	83

The IRAP data from Experiment 1 were in accordance with the pattern of preferences predicted by perceived social similarity—participants responded more quickly when confirming preferences for Irish over Scottish, Scottish over American, and American over African than when confirming the opposite pattern of preferences. Interestingly, the preferences obtained on the Likert scales diverged from the IRAP in two ways: No differential preference was shown between Irish and Scottish, and Africans were preferred to Americans.

Experiment 2

Although the pattern of preferences obtained from the IRAP in Experiment 1 was in accordance with predictions, this was the first study to produce such an effect, and it was obtained with only one specific sample of participants. In Experiment 2, therefore, we made another set of predictions, again based on perceived social similarity, but using a non-Irish sample. Pilot work had shown that Irish American individuals, given a choice among Irish, Scottish, and African, typically rated Irish as most similar to American, with Scottish ranked next, followed by African. Based on this ranking it was predicted that Irish American participants might show a preference for American over Irish, Irish over Scottish, and Scottish over African. An IRAP and Likert measures, similar to those employed in Experiment 1, were used to test these predictions.

Method

Participants

Sixteen individuals, 5 male and 11 female, who were citizens of the United States of America, and who self-identified as “Irish American,” agreed to participate in Experiment 2. Ages ranged from 18 to 35 years ($M = 23$ years). Eight participants each were assigned randomly to the consistent-relations-first or inconsistent-relations-first condition.

Setting

The setting used in Experiment 1 was also employed in Experiment 2.

Apparatus and Materials

The four Likert scales employed in Experiment 1 were also employed in Experiment 2. The IRAP software employed in Experiment 1 was used in Experiment 2, but the target stimuli were different. Specifically, the six target word pairs presented were *American-Irish*, *Irish-Scottish* and *Scottish-African*, and their reversed counterparts (see Table 5).

Table 5
IRAP Stimuli in Experiment 2

Sample 1 More Likeable	Sample 2 Less Likeable
Response option 1 True	Response option 2 False
Targets deemed consistent with Sample 1	Targets deemed consistent with Sample 2
American Irish Irish Scottish Scottish African	Irish American Scottish Irish African Scottish

Procedure

The procedure for Experiment 2 was identical to that of Experiment 1.

Results

Explicit Likert Analyses

A mean Likert rating for each of the four social groups was calculated across the 16 participants (American, $M = 3.5$, $SE = .376$; Irish, $M = 4.438$, $SE = .223$; Scottish, $M = 3.313$, $SE = .445$; African, $M = 3.188$, $SE = .277$). Irish was rated as more likeable than any other group, with American, Scottish, and African rated as almost equally likeable. The four ratings from each participant were entered into a one-way repeated measures ANOVA, which yielded a significant result, $F(15, 3) = 4.071$, $p = .0122$, $\eta_p^2 = .2$. Post hoc Fisher's PLSD tests indicated that the ratings for Irish were significantly different from the ratings for every other social group ($ps < .007$). All other comparisons were nonsignificant ($ps > .4$). This pattern of preferences was not consistent with the prediction made in terms of perceived group similarity, in that (a) Irish American participants did not rate American as more likeable than Irish, and (b) no significant differences were found among the likeability ratings for American, Scottish, and African.

IRAP Preliminary Analyses

Greenwald et al.'s (2003) adapted C4 algorithm was again applied to the IRAP data. The same preliminary analyses were conducted as in Experiment 1. The overall means and standard errors calculated across participants are presented in Table 6. Once again, the mean response latencies, within each pair of test blocks, were consistently shorter for the consistent relative to the inconsistent trials.

Table 6
Overall Adjusted Mean Latencies and Standard Errors in Each Condition for Each Test Block

Test order	IRAP test type per block											
	Consistent						Inconsistent					
	Test 1		Test 2		Test 3		Test 1		Test 2		Test 3	
	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>	<i>M</i>	<i>SE</i>
Consistent first	2,395	60	2,326	62	2,344	57	2,602	52	2,441	43	2,457	46
Inconsistent first	2,214	71	2,183	91	2,148	54	2,397	65	2,370	77	2,236	85

A $2 \times 2 \times 3$ mixed repeated measures ANOVA indicated that the difference in latencies between consistent and inconsistent trials reached statistical significance, $F(1, 14) = 22.068$, $p = .0003$, $\eta_p^2 = .6$. The main effect for test exposure also proved to be significant, $F(2, 14) = 23.068$, $p < .001$, $\eta_p^2 = .6$, with overall latencies reducing across exposures. Furthermore, exposure interacted significantly with test order, $F(2, 28) = 5.947$, $p < .007$, $\eta_p^2 = .3$, indicating that the differences in overall latencies between consistent- and inconsistent-first conditions differed across the three test exposures. Given that neither of these results was directly relevant to the critical IRAP effect, they were not subject to further analyses.

Although IRAP condition did not interact significantly with any of the other variables, the interaction with test exposure did approach significance, $F(2, 14) = 3.256, p = .054, \eta_p^2 = .2$. Three follow-up repeated measures ANOVAs indicated that the predicted difference between consistent and inconsistent blocks remained significant ($ps < .02$) for each pair of test blocks, although the effect size was halved for the latter two pairs of test blocks relative to the first ($\eta_p^2 = .8$ vs. $.4$). Given that the IRAP effect was significant and in the predicted direction for each pair of test blocks, and the interaction only approached significance and had not been observed in Experiment 1, test exposure was removed from subsequent analyses.

Individual Pairs Analyses

The mean response latencies for each participant for consistent and inconsistent trials were calculated for each of the three target pairs (see Table 7). For each pair, response latencies were shorter for consistent relative to inconsistent trials. A 2×3 ANOVA was conducted with test type (consistent vs. inconsistent) and target pair (1–3) as repeated measures. Ignoring the previously reported main effect for test type, the effect for target pair was significant, $F(2, 30) = 23.788, p < .0001, \eta_p^2 = .6$. Three separate ANOVAs were used to analyze consistent versus inconsistent latency differences for each pair. In each case, the difference proved to be significant (all $ps < .008$). Overall, therefore, each target pair produced the predicted IRAP effect.

Table 7
Overall Mean Latencies and Standard Errors for Consistent and Inconsistent Trials for Each Target Pair in Experiment 2

Target pair	IRAP test type	<i>M</i>	<i>SE</i>
American-Irish	Consistent	2,118	61
	Inconsistent	2,257	59
Irish-Scottish	Consistent	2,340	56
	Inconsistent	2,519	33
Scottish-African	Consistent	2,346	58
	Inconsistent	2,463	61

Explicit-Implicit Correlations

For the IRAP data, each participant's mean latency score for consistent trials was subtracted from that for inconsistent trials for each pair. For the Likert data, each participant's rating for Irish was subtracted from American, and, similarly, Scottish from Irish and African from Scottish. All of the three Spearman rank correlations were nonsignificant, although the effects for the American-Irish and Irish-Scottish pairs were negative (American-Irish, $r_s = -.209, p = .42$; Irish-Scottish, $r_s = -.19, p = .47$; and Scottish-African, $r_s = .271, p = .29$).

Trial-Type Analyses

Similar to Experiment 1, the trial-type data were subjected to a $2 \times 2 \times 2$ ANOVA (see Table 8), with sample, target type, and consistent versus

inconsistent as repeated measures. The effect for IRAP test type (consistent vs. inconsistent) did not interact with either of the other two variables (all $ps > .2$). Also similar to Experiment 1, a main effect for sample was recorded, $F(1, 15) = 54.246$, $p < .0001$, $\eta_p^2 = .8$, indicating that participants responded more quickly on "More Likeable" than "Less Likeable" trials ($M = 2,267$ ms, $SE = 34$, and $M = 2,420$ ms, $SE = 27$, respectively).

Table 8
Overall Mean Latencies and Standard Errors for Each Trial Type for Consistent and Inconsistent Trials

Sample	Target	IRAP test type	<i>M</i>	<i>SE</i>
Less Likeable	Less Likeable to More Likeable	Consistent	2,305	73
Less Likeable	Less Likeable to More Likeable	Inconsistent	2,466	45
Less Likeable	More Likeable to Less Likeable	Consistent	2,415	47
Less Likeable	More Likeable to Less Likeable	Inconsistent	2,491	36
More Likeable	Less Likeable to More Likeable	Consistent	2,187	62
More Likeable	Less Likeable to More Likeable	Inconsistent	2,367	72
More Likeable	More Likeable to Less Likeable	Consistent	2,158	63
More Likeable	More Likeable to Less Likeable	Inconsistent	2,356	60

Conclusions

The IRAP data from Experiment 2 were again in accordance with the pattern of preferences predicted by perceived social similarity, thus diverging from the preferences obtained on the Likert scales.

General Discussion

The IRAP data from Experiments 1 and 2 were in accordance with the pattern of preferences predicted by perceived social similarity. In comparing the ratings obtained from the Likert scales with the IRAP results, it is clear that the two measures diverged. The pattern of preferences obtained in Experiment 1 from the Likert measures indicated that participants did not discriminate between Irish and Scottish; and, furthermore, African was rated as significantly more likeable than American (the IRAP showed a bias for Irish over Scottish and American over African). In Experiment 2, the pattern of preferences obtained from the Likert measures indicated that Irish was rated as significantly more likeable than any other group; and, furthermore, participants did not discriminate between American, Scottish, and African (the IRAP showed an American-over-Irish-over-Scottish-over-African pattern of biases). Finally, all but one of the correlational analyses between the IRAP and Likert measures were nonsignificant. Furthermore, the divergence between the two measures occurred even though one measure was presented immediately after the other, thus allowing for possible carry-over effects. The current findings therefore provide an example of the classic separation between an explicit and an implicit measure of a socially or politically sensitive attitude or belief (De Houwer, 2002, 2006; Nosek et al., 2006).

As explained in the Introduction, the predictions for the IRAP performances were based on the structural similarity between the IAT and IRAP. The current findings therefore support the conclusion that the IRAP effect is based, at least to some degree, on manipulating relevant S-R compatibility. In drawing this conclusion, however, it seems important to propose a possible explanation, at the level of psychological process, of the IRAP effect itself. Barnes-Holmes et al. (2006) provided one possible explanation as follows. Each trial of the IRAP presents a target stimulus with contextual cues that specify particular relational (e.g., *same* vs. *opposite*) and functional (e.g. *pleasant* vs. *unpleasant*) dimensions, which produces an immediate relational response before the participant actually presses a response key (the participant may or may not be consciously aware of this response). It is assumed that the probability of the initial response will be determined by the verbal and nonverbal history of the participant and current contextual variables. By definition, the most probable immediate response will be emitted first most often, and thus, during a consistent IRAP trial, that response will tend to match the correct key; during an inconsistent trial, however, the probability of such a match will be reduced. Thus, across multiple trials, the average latency for inconsistent blocks will be longer than for consistent blocks. In short, the IRAP effect is based on immediate and perhaps unconscious relational responding, which is made apparent to the researcher when the behavioral system is put under pressure to respond quickly and accurately.

Why did the IRAP and Likert measures diverge in the current study? Although any response to this question must remain quite speculative, it seems likely that responses to the Likert scales were influenced to some degree by variables that extended beyond the relational or contextual control of likeability *per se*. For example, some participants in Experiment 1 may have rated "African" more likeable than "American" because doing so cohered with other instances of their relational responding. Imagine, for example, that a participant's deictic relations involved extended relational networks, such as, "I think America exploits Third World countries," and actions that are generally coordinated with such networks. When presented with the Likert scales, responding in a way that coheres with such networks is relatively straightforward (do not provide a rating that favors the "super power" over poor and oppressed nations). When exposed to the IRAP, however, the impact of the participant's "politically correct" deictic relational responding would be much reduced because there is insufficient time, on a trial-by-trial basis, to engage in the additional and sometimes complex relational activity that serves to generate a "carefully considered" answer. Broadly similar explanations may be offered for other instances in which the IRAP and Likert measures diverged. For example, Irish and Scottish are generally perceived to be very similar, and thus it is possible that some participants did not rate the former as more likeable than the latter because doing so might be seen as excessively parochial or patriotic. In this context, it is also worth noting that in Experiment 2 the Irish American participants rated the Irish as more likeable than the other three groups, but they provided these ratings in an Irish University setting to an Irish experimenter. In short, therefore, the IRAP effect may have diverged from the explicit measure because the rapid-response requirement significantly reduced the impact of those extended relational networks that may have "contaminated" the targeted relational responding (cf. Wilson,

Lindsey, & Schooler, 2000). Additional research will be needed to subject this interpretation to empirical analysis, but given the clear IRAP effect obtained in the current study, such work would appear to be worthwhile.

In conducting future research on the IRAP, it will be important to determine to what extent the IRAP may be considered a measure of implicit cognition. De Houwer (2006) argued recently that a measure is implicit if one or more of the following criteria apply: Participants (a) are not aware that the relevant attitude is being measured, (b) do not have conscious access to the attitude, or (c) have no control over the outcome of the measure. In the context of the current study, participants (in Experiments 1 and 2) were almost certainly aware that their "attitudes" toward the various social groups were being assessed in some manner. With respect to the second criterion, the extent to which the participants were actually aware of their relative preferences remains uncertain. That the IRAP effects often diverged from the explicit ratings could be taken to indicate that the participants were not aware. On balance, it is possible that participants were fully aware and simply lied about their preferences on the explicit measures (on the grounds of political correctness). Even if this was the case, however, the fact that the participants failed to produce politically correct responses on the IRAP indicates that the measure meets the third criterion (i.e., they could not control the measurement outcome). Indeed, recent findings from our research group indicate that participants possess limited control over the IRAP effect even when they are told how the IRAP works and are expressly instructed to produce a "faked" performance (McKenna, Barnes-Holmes, Barnes-Holmes, & Stewart, 2007). Overall, therefore, current and previous findings show that the IRAP *might* be defined as an implicit measure according to criterion 2, and more certainly according to criterion 3.

Assuming that the IRAP may be used as an implicit measure, as defined in terms of either criterion 2 or criterion 3, it seems important to ask, What exactly does the IRAP provide above and beyond the IAT (and, indeed, above other currently available implicit measures)? De Houwer (2002) has argued that the IAT measures associations rather than relations among stimuli or events, and as such can provide only an indirect measure of beliefs:

Greenwald et al. (1998) designed the IAT to assess the strength of associations between concepts in memory. One can argue that beliefs involve more than just associations between concepts. First, beliefs reflect qualified associations. For instance, the belief "I am a bad person" implies *a special type of association* [italics added] between the concept "self" and the concept "bad," namely *a directional association* [italics added] which specifies that "bad" is a property or characteristic of "self." IAT effects do not reflect the *nature or directionality of an association between concepts* [italics added], they can reflect only strength of association. Second, many beliefs involve several associations and several concepts. For instance, conditional beliefs such as "if I do not perform well on a task, then I am an inferior person" involve rather complex structures of qualified associations between several concepts. The IAT cannot be used to directly capture such complex conditional beliefs (also see de Jong et al.,

2001, p. 111). . . . In sum, the IAT does not provide a measure of beliefs, nor was it designed to do so. It can only provide an index of associations that are assumed to be involved in certain beliefs and thus indirect evidence for the presence of certain beliefs. (pp. 117-118).

A key limitation to the IAT, therefore, is that it cannot readily assess the nature or directionality of an association (i.e., a relation), and, furthermore, it cannot be used to assess a complex structure of directional associations (i.e., a relational network). Indeed, the same criticism may be leveled at any implicit measure that relies on the putative measurement of associative strength, such as evaluative priming (Fazio, Sanbonmatsu, Powell, & Kardes, 1986), the Extrinsic Affective Simon Task (EAST; De Houwer, 2003b), and the Go/No Go Association Task (GNAT; Nosek & Banaji, 2001). In contrast to mainstream measures, the IRAP targets specific relations among sets of stimuli by asking participants to respond directly to target statements concerning the actual attitudes or beliefs under study (e.g., are Irish more likeable than Scottish?). Consequently, an IRAP effect may provide a more direct measure, requiring less inference on behalf of the researcher, than a procedure that taps into the associations that are assumed to underlie a particular attitude or belief. If the IRAP does indeed prove to be a relatively direct measure, it would provide an important supplement to tests of implicit cognition that rely on associative assumptions.

Contrasting the IRAP with other implicit measures serves to highlight the way in which Relational Frame Theory differs in its approach to implicit cognition relative to mainstream psychology. From an RFT perspective, the concept of implicit cognition refers to probabilistic patterns of relational responding that are revealed when “contaminating” sources of contextual control are removed, or at least reduced, as is the case when responding under high time-pressure is required. According to mainstream psychology, implicit measures are often seen as tapping into “underlying” associative networks, which, by definition, are not directional in nature (if “Bad” is associated with “Insect,” “Insect” is associated with “Bad”). Indeed, according to some mainstream researchers, implicit cognition is almost completely based on these types of “simple” associations (Gawronski, LeBel, & Peters, 2007). According to RFT, however, a variety of stimulus relations, or what De Houwer (2006) calls “directional associations,” may also be involved in implicit cognition, and the current IRAP data appear to support this latter view. It should be recognized, however, that considerable empirical work will be required to explore more fully the implications of the RFT view of implicit cognition and how it compares and contrasts with mainstream views.

Clearly, many issues remain to be addressed in the study of implicit cognition and in the development and refinement of procedures designed to reveal implicit beliefs and attitudes. For example, it has yet to be determined if performance on an IRAP will predict “prejudiced” behavior in the natural environment, and thus caution is required in concluding that implicit prejudice necessarily leads to discriminatory action. Nevertheless, the current study suggests that the IRAP may be used to reveal politically or socially sensitive attitudes, when comparing particular national groups, which might not be expressed using a typical explicit measure.

References

- BARNES-HOLMES, D., HEALY, O., & HAYES S. C. (2000). Relational frame theory and the relational evaluation procedure: Approaching human language as derived relational responding. In J. C. Leslie & D. E. Blackman (Eds.), *Experimental and applied analyses of human behaviour* (pp. 149–180). Reno, NV: Context Press.
- 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 Evaluation Procedure (IRAP) as a direct measure of implicit beliefs. *The Irish Psychologist*, *32*(7), 169–177.
- CATANIA, A. C. (1998). *Learning* (4th ed.). Englewood Cliffs, NJ: Prentice Hall.
- COTA, A. A., EVANS, C. R., DION, K. L., KILIK, L., & LONGMAN, R. S. (1995). The structure of group cohesion. *Personality and Social Psychology Bulletin*, *21*, 572–580.
- CULLINAN, V., BARNES, D., & SMEETS, P. M. (1998). A precursor to the relational evaluation procedures: Analyzing stimulus equivalence. *The Psychological Record*, *48*, 121–145.
- CULLINAN, V., BARNES-HOLMES, D., & SMEETS, P. M. (2000). A precursor to the relational evaluation procedure: Analyzing stimulus equivalence II. *The Psychological Record*, *50*, 467–492.
- CULLINAN, V., BARNES-HOLMES, D., & SMEETS, P. M. (2001). A precursor to the relational evaluation procedure: The search for the contextual cues that control equivalence responding. *Journal of the Experimental Analysis of Behavior*, *76*, 339–349.
- DE HOUWER, J. (2002). The Implicit Association Test as a tool for studying dysfunctional associations in psychopathology: Strengths and limitations. *Journal of Behaviour Therapy and Experimental Psychiatry*, *33*, 115–133.
- DE HOUWER, J. (2003a). A structural analysis of indirect measures of attitudes. In J. Musch & K. C. Klauer (Eds.), *The psychology of evaluation: Affective processes in cognition and emotion* (pp. 219–244). London, England: Erlbaum Associates.
- DE HOUWER, J. (2003b). The extrinsic affective Simon task. *Experimental Psychology*, *50*, 77–85.
- 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.
- FAZIO, R. H., SANBONMATSU, D. M., POWELL, M. C., & KARDES, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, *50*, 229–238.
- GAWRONSKI, B., LEBEL, E. P., & PETERS, K. R. (2007). What do implicit measures tell us? Scrutinizing the validity of three common assumptions. *Perspectives on Psychological Science*, *2*(2), 181–193.
- 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.

- GREENWALD, A. G., BANAJI, M. R., RUDMAN, L. A., FARNHAM, S. D., NOSEK, B. A., & MELLOTT, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review*, *109*, 3-25.
- 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.
- HAYES, S. C., BARNES-HOLMES, D., & ROCHE, B. (2001). *Relational frame theory: A post-Skinnerian account of human language and cognition*. New York: Plenum.
- MCKENNA, I., BARNES-HOLMES, D., BARNES-HOLMES, Y., & STEWART, I. (2007). Testing the fake-ability of the Implicit Relational Assessment Procedure (IRAP): The first study. *International Journal of Psychology and Psychological Therapy*, *7*(2), 253-268.
- NOSEK, B. A., & BANAJI, M. R. (2001). The go/no-go association task. *Social Cognition*, *19*, 625-666.
- NOSEK, B. A., GREENWALD, A. G., & BANAJI, M. R. (2006). The Implicit Association Test at age 7: A methodological and conceptual review. In J. A. Bargh (Ed.), *Automatic processes in social thinking and behavior* (pp. 265-292). Hove, UK: Psychology Press.
- O'HORA, D., BARNES-HOLMES, D., ROCHE, B., & SMEETS, P. M. (2004). Derived relational networks and control by novel instructions: A possible model of generative verbal responding. *The Psychological Record*, *54*, 437-460.
- O'HORA, D., PELAEZ, M., & BARNES-HOLMES, D. (2005). Derived relational responding and performance on verbal sub-tests of the WAIS-III. *The Psychological Record*, *55*, 155-175.
- STEWART, I., BARNES-HOLMES, D., & ROCHE, B. (2002). Developing an ecologically valid model of analogy using the relational evaluation procedure. *Experimental Analysis of Human Behavior Bulletin*, *20*, 12-16.
- STEWART, I., BARNES-HOLMES, D., & ROCHE, B. (2004). A functional-analytic model of analogy using the relational evaluation procedure. *The Psychological Record*, *54*, 531-552.
- TAJFEL, H. (1982). Social psychology of intergroup relations. *Annual Review of Psychology*, *33*, 1-39.
- WILSON, T. D., LINDSEY, S., & SCHOOLER, T. Y. (2000). A model of dual attitudes. *Psychological Review*, *107*, 101-126.

Journal of the Experimental Analysis of Behavior

EDITOR

James E. Mazur

EXECUTIVE EDITOR

Victor G. Laties

ASSOCIATE EDITORS

Brent Alsop

Karen M. Lionello-DeNolf

Anthony P. McLean

Amy L. Odum

EDITOR FOR**BEHAVIORAL
PHARMACOLOGY**

Jonathan L. Katz

EDITOR FOR**BEHAVIORAL
NEUROSCIENCE**

David W. Schaal

REVIEW EDITOR

M. Jackson Marr

MANAGING EDITOR

Sharon Corcoran

BUSINESS MANAGER

Monica Bonner

EDITORIAL ASSISTANT

Marcus Jeffries

BOARD OF EDITORS

Peter Balsam

Terry W. Belke

Matthew C. Bell

Jack Bergman

Christopher M. Bradshaw

Philip N. Chase

Jesse Dallery

William V. Dube

Edmund Fantino

J. Gregor Fetterman

Lanny Fields

Stephen C. Fowler

Timothy D. Hackenberg

David N. Harper

Jennifer J. Higa

Christine E. Hughes

Iver H. Iversen

Eric A. Jacobs

Peter Killeen

Jason Landon

Armando Machado

Suzanne H. Mitchell

Michael A. Nader

Allen Neuringer

Michael Perone

Cynthia J. Pietras

Raymond C. Pitts

Jennifer L. Raymond

Ruth Anne Rehfeldt

Robert A. Rescorla

Susan M. Schneider

Julie B. Schweitzer

Timothy A. Shahan

Richard L. Shull

William Timberlake

François Tonneau

Peter J. Urciuoli

John H. Wearden

Gail Winger

Alice M. Young

Thomas R. Zentall

Website: seab.envmed.rochester.edu/jeab E-mail: jeab@indiana.edu

Consider a subscription to the journal for original research relevant to the behavior of individual organisms. Review articles and theoretical papers also considered for publication. Six issues published yearly by the Society for the Experimental Analysis of Behavior.

– JULY 2009 –

N. Will Shead and David C. Hodgins. Probability discounting of gains and losses: Implications for risk attitudes and impulsivity

Celia Lie and Brent Alsop. Effects of point-loss punishers on human signal-detection performance

Christopher E. Bullock and Todd M. Myers. Stimulus-food pairings produce stimulus-directed touch screen responding in cynomolgus monkeys (*Macaca fascicularis*) with or without a positive response contingency

Lanny Fields, Erica Doran, and Michael Marroquin. Equivalence class formation in a trace stimulus pairing two-response format: Effects of response labels and prior programmed transitivity induction

Miguel Rodríguez Valverde, Carmen Luciano, and Dermot Barnes-Holmes. Transfer of aversive respondent elicitation in accordance with equivalence relations

– TECHNICAL ARTICLE –

Eunjeong Lee, Ana I. Oliveira-Ferreira, Ed de Water, Hans Gerritsen, Mattijs C. Bakker, Jan A.W. Kalwij, Tjerk van Goudoever, Wietze H. Buster and Cyriel M. A. Pennartz. Ensemble recordings in awake rats: achieving behavioral regularity during multimodal stimulus processing and discriminative learning

– BOOK REVIEW –

Howard Rachlin and Marvin Frankel. Taking pragmatism seriously: A review of William Baum's *Understanding behaviorism: Behavior, culture, and evolution (Second Edition)*

– SEPTEMBER 2009 –

Greg Jensen and Allen Neuringer. Barycentric extension of generalized matching

Erin B. Rasmussen and M. Christopher Newland. Quantification of ethanol's anti-punishment effect in humans using the generalized matching equation

Matthew J. Pizzo, Kimberly Kirkpatrick, and Pamela J. Blundell. The effect of changes in criterion value on differential reinforcement of low rate schedule performance

Christopher A. Podlesnik, Corina Jimenez-Gomez, Ryan D. Ward, and Timothy A. Shahan. Resistance to change and frequency of response-dependent stimuli uncorrelated with reinforcement

Rogelio Escobar and Carlos A. Bruner. Observing responses and serial stimuli: Searching for the reinforcing properties of the S-

Paula Debert, Edson M. Huziwara, Robson Brino Faggiani, Maria Eugênia Simões de Mathis, and William J. McIlvane. Emergent conditional relations in a go/no-go procedure: Figure-ground and stimulus-position compound relations

Barry Haimson, Krista M. Wilkinson, Celia Rosenquist, Carolyn Ouimet, & William J. McIlvane. Electrophysiological correlates of stimulus equivalence processes

Troy J. Zarcone, Rong Chen, and Stephen C. Fowler. Effects of differing response-force requirements on food-maintained responding in C57BL/6j mice

– BOOK REVIEW –

Edward K. Morris. Behavior analysis and ecological psychology: Past, present, and future: A review of Harry Heft's *Ecological Psychology in Context*

JEAB, Monica A. Bonner

Department of Psychological & Brain Sciences

Indiana University

1101 E. Tenth Street

Bloomington, IN 47405-7007 USA

American Express, VISA, and MasterCard orders are accepted.