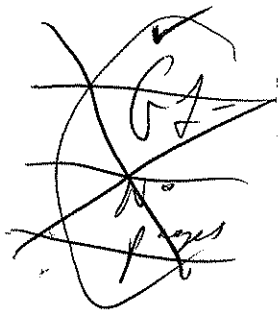


04 F exp #3



p3 24

- (10) DOMAIN: A/1= race, B/2= ethnicity C/3 = gender-sex D/4 = food or drink  
E/5 = other consumer F/6= political G/7 = drugs or tobacco  
H/8 = self esteem J/9= personality/self K/10 = clinical  
L/11= relationships M/12= other? (not a tony category)

p3 26  
p3 25

- (11) BEHAVIOR: single=1, average=2  
(12) IAT TYPE: attitude=1, belief=2, self=3, not reported = 4

- p3 26 (13) EM TYPE: attitude=1, belief=2, self=3, not reported = 4

- p3 26 (14) OVERALL METHOD: not=0, observed=1

- p3 26 (15) METHOD: RepPast=1, future=2, emotion=3, judge=4, obs=5, neuro=6, other=7

- p3 28 (16) SCORE: millisecond=0, log=1, algorithm=2, NotReported=3

- p3 25 (17) words=0, pictures=1, NotReported=2

- p3 25 (18) number of IATs: 2 each

- (19) IAT ORDER: NotReported=0, iatfirst=1, iatsecond=2, iatthird=3

p3 25  
p3 26

- (20) EXPLICIT ORDER: NotReported=0, explicitfirst=1, expsecond=2, explthird=3

- (21) BEHAVIOR ORDER: NotReported=0, behfirst=1, behsecond=2, behthird=3

- (22) IAT vs. behavior: NotReported=0, before=1, after=2, counter=3

- (23) EXPLICIT vs. beh: NotReported=0, explicitfirst=1, expsecond=2, counter=3

- (24) IAT SESSION: same=0, different=1

- (25) EXPLICIT SESSION: same=0, different=1

- p3 25 (26) IAT SOCIAL DESIRABILITY 1-7 1.5

- p3 26 (27) EXPLICIT SOCIAL DESIRABILITY 1-7 1.5

- p3 26 (28) BEHAVIOR CONTROLLABLE: 1-10 10

- p3 25, 26 (29) IAT SPECIFIC 1-7 5.5

- p3 26 (30) EXPLICIT SPECIFIC 1-7 5.5

NONE

- p3 25 (31) OPPOSITION 1-5 1.5

- p3 25 (32) RACIAL 0=not, 1=racial

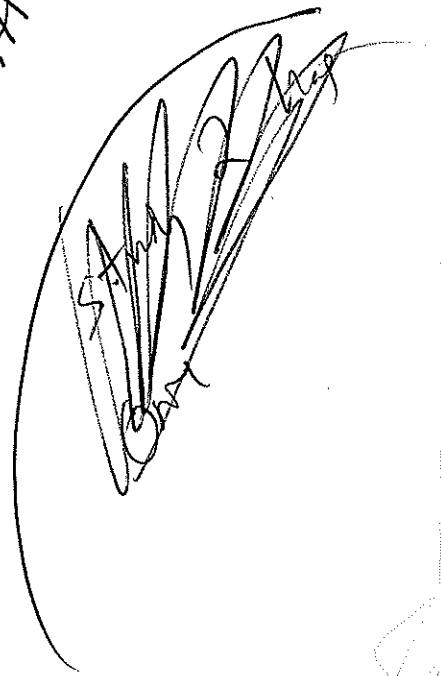
- p3 25 (33) type of iat: single=1, dual=2, personalized=3

1/2 1/2

No ELCs reported in review

Also, separate?

together



O & F #4

67-  
No  
pages

(10) DOMAIN: A/1= race, B/2= ethnicity C/3 = gender-sex D/4 = food or drink

E/5 = other consumer F/6= political G/7 = drugs or tobacco

H/8 = self esteem I/9= personality/self K/10 = clinical

L/11= relationships M/12= other? (not a tony category)

p1 33 (11) BEHAVIOR: single=1, average=2

p1 32 (12) IAT TYPE: attitude=1, belief=2, self=3, not reported = 4

p1 33 (13) EM TYPE: attitude=1, belief=2, self=3, not reported = 4

p1 33 (14) OVERALL METHOD: not=0, observed=1

(15) METHOD: RepPast=1, future=2, emotion=3, judge=4, obs=5, neuro=6, other=7

p1 33 (16) SCORE: millisecond=0, log=1, algorithm=2, NotReported=3

(17) words=0 pictures=1, NotReported=2

(18) number of IATs: 1 (per subject)

(19) IAT ORDER: NotReported=0, iatfirst=1, iatsecond=2, iatthird=3

(20) EXPLICIT ORDER: NotReported=0, explicitfirst=1, expsecond=2, explthird=3

(21) BEHAVIOR ORDER: NotReported=0, behfirst=1, behsecond=2, behthird=3

(22) IAT vs. behavior: NotReported=0, before=1, after=2, counter=3

(23) EXPLICIT vs. beh: NotReported=0, explicitfirst=1, expsecond=2, counter=3

(24) IAT SESSION: same=0, different=1

(25) EXPLICIT SESSION: same=0, different=1

(26) IAT SOCIAL DESIRABILITY 1-7 2

(27) EXPLICIT SOCIAL DESIRABILITY 1-7 2

(28) BEHAVIOR CONTROLLABLE: 1-10 9

(29) IAT SPECIFIC 1-7 5

(30) EXPLICIT SPECIFIC 1-7 5

(31) OPPOSITION 1-5 5

(32) RACIAL: 0=not, 1=racial

(33) type of iat: single=1, dual=2, personalized=3

2 for some diff  
but could  
partly affect

by default

zero - 5  
but

1/2

appropriate. Although it is certainly socially desirable to present oneself as healthy, and hence, individuals may be more likely to claim to like apples relative to candy bars, they are probably somewhat less motivated to behave in a socially desirable fashion here than when completing a direct, explicit measure of racial attitudes. Recall that Karpinski and Hilton (2001) observed null relations between an IAT assessing preferences for apples versus candy bars and explicit measures of attitudes toward the same objects. If the personalized version of the IAT removes some of the contamination of extra-personal associations, then we should expect it to correlate better with explicit measures and behavioral intentions in this relatively less socially sensitive domain. We test this hypothesis in Experiment 3.

### Experiment 3

We have argued that Blacks are portrayed relatively negatively by society, and that this information can be used in a way that makes people appear relatively prejudiced on the IAT. Similarly, Karpinski & Hilton (2001) reported an "apples-candy bar" IAT where participants appeared far more positive toward apples than explicit measures and actual choice behavior indicated. In Experiment 3, we again tested the extra-personal association hypothesis, but because of apples' relatively positive portrayal, we predicted that a traditional IAT would show positivity toward apples relative to candy bars (thus replicating Karpinski & Hilton's findings). However, we predict that this positivity will be less apparent on a personalized IAT. And although participants might still be motivated to present themselves in a socially desirable light by claiming to engage in healthy eating habits, the inclusion of explicit measures in this less socially sensitive domain allows us to test the prediction that a personalized IAT will correlate better than a traditional IAT with explicit measures of attitudes, past behavior, and behavioral intentions.

10

Participants Sixty-two undergraduates at a Midwestern university participated for course credit. Three were omitted from analyses for committing a large number ( $> 20\%$ ) of errors or because of missing data, resulting in a final sample of 35 women and 24 men.

Materials and Procedure. Participants were seated in individual cubicles and read a set of instructions that described the IAT as being about "categorization skills." They were randomly assigned to either the traditional or personalized IAT condition (IAT type was manipulated just as in Experiment 1, which included changes in both normatively pleasant and unpleasant items and category labels). Parameters of the IAT were modeled after Karpinski & Hilton (2001), with some minor exceptions noted below. Participants were told that they would be categorizing a variety of different items, that instructions on the screen would describe to them how to categorize the items, and to press any key to begin. There were 7 blocks in the IAT, each with 50 trials. The pleasant/unpleasant and liked/disliked items were the same as those used in the first experiment. The apple and candy-bar related items consisted of words related to the 2 categories (e.g., "Snickers," "Red Delicious"). Some of these items were taken from Karpinski & Hilton (2001), and others were derived from our own pre-testing (the complete list is presented in the Appendix). The first two blocks consisted of practice with the categorization first of candy-bar and apple related items, and then pleasant and unpleasant items, respectively. Blocks 3 and 4 were critical combined blocks, where candy-bar-related items were associated with the positive category, and apple-related items were associated with the negative (or vice versa, depending on the counter-balancing conditions to which participants were assigned). Block 5 was a practice block consisting of candy-bar and apple-related items. Blocks 6 and 7 were also critical combined blocks, and were identical to blocks 3 and 4, but the food that was

IAT

17

19-25

19-25

associated with the positive category was now associated with the negative (and the food that was associated with the negative category was now associated with positive).

After completing the IAT, participants completed several explicit measures of their attitudes toward apples and candy bars, which were introduced as "measures of certain beliefs that might affect the categorization skills in which we were interested" (some of which were taken from Karpinski & Hilton, 2001). These included several semantic differential items (Ugly/Beautiful, Bad/Good, Unpleasant/Pleasant, Foolish/Wise, and Awful/Nice), liking ("How much do you like eating apples [candy bars]?"), a behavioral measure ("Do you eat apples [candy bars] often?"), and a forced choice measure of behavioral intention ("If given a choice between an apple and a candy bar, which would you choose?"), all using 7-point scales. Next, they completed a feeling thermometer (on a 0 – 100 scale) regarding the extent of their favorability toward several filler foods along with our foods of interest. Finally, participants provided rank order information on their preferences for these foods. They were then debriefed, thanked, and dismissed.

## Results and Discussion

Effects of IAT version. Critical block means were derived as in Experiments 1 and 2, and were submitted to a 2 (Block Type: Apple/+ v. Apple/-) X 2 (IAT Type: Traditional v. Personalized) ANOVA, with repeated measures on the first factor. Only a marginal Block Type X IAT Type interaction emerged,  $F(1, 57) = 2.90, p = .09$ . On the traditional IAT, the mean of response latencies on the Apple/- task was 834.81 ( $SD = 120.30$ ), compared to 780.23 ( $SD = 141.80$ ) on the Apple/+ task. On the personalized IAT, the Apple/- task mean was 947.50 ( $SD = 200.71$ ), compared to 950.89 ( $SD = 197.28$ ) on the Apple/+ task. Accuracy on the critical apple-candy bar trials did not differ as a function of IAT version (96% for each,  $t < 1$ ). Given that the

Behavior

28

29

30

EMS

27

19  
25

13

14

15

11

The explicit measures all correlated significantly with one another. With the exception of the semantic differential, whose correspondence with the other measures was moderate ( $r$ 's from .44 to .62), correlations between explicit measures were quite high ( $r$ 's from .57 to .87). Hence, a single index of explicitly measured attitudes was derived as the average of the standardized individual measures ( $\alpha = .91$ ). As indicated in Table 2, this composite index also correlated strongly with the personalized IAT, and only non-significantly with the traditional IAT. Tests of the difference between the two IATs' correlations with the explicit measures revealed statistically significant differences for several of the individual measures, as well as the overall composite measure. Table 2 provides the results of these tests for each measure.

Recall that in Experiments 1 and 2, IAT scores were also calculated using the algorithm prescribed by Greenwald et al. (2003), except that we did not implement an error penalty and we used transformed latencies. Experiment 3 was modeled after a number of IAT studies more recent than the original Greenwald et al. (1998) work. Hence, the length of the task had been reduced to 20 practice and 40 critical trials for each of the combined task blocks, just as was true for the web-based IATs examined by Greenwald et al. (2003). Their scoring algorithm dictates that the  $D$  score (mean difference divided by the pooled standard deviation) be computed separately for the practice block and for the critical block, and that these two scores then be averaged as the IAT index. However, given the differing lengths of the practice and critical blocks, this practice means that trials on the practice blocks are given twice the weight of those on the critical blocks. We view assigning this disproportionate weight to practice trials to be unwarranted, so in adopting the algorithm to Experiment 3, we computed difference scores and standard deviations such that all trials received equal weight (i.e., means and standard deviations were computed across all 60 trials). Using this algorithm, the traditional IAT ( $M = .17$ ,  $SD =$

16

Participants. Individuals were recruited from student newspaper advertisements and flyers posted on the campus of a Midwestern university for participation in this and other unrelated experiments in exchange for 20 dollars during the summer of 2003. Forty-nine individuals, all of whom had listed hometowns within the United States on a preliminary background questionnaire, served as the participants. Of these, data from one participant were omitted because of high errors on the IAT (>20%), and from another because of equipment failure, resulting in 18 male and 29 female participants.

Materials and Procedure. Instructions and procedures were analogous to those of Experiment 3, but parameters and stimuli for the IAT were modeled after Greenwald et al. (2003). The pleasant/unpleasant items for both versions of the IAT were normatively pleasant and unpleasant, and the Bush and Gore items consisted of their full names and last names only in Black capital letters, and 2 head-shot photos each of Bush and Gore. There were 7 blocks in total, with practice blocks consisting of 20 trials each and critical blocks consisting of 40 trials each. Blocks 1 and 2 consisted of practice categorizing first Bush and Gore items, and pleasant and unpleasant items, respectively. Blocks 3 was a practice combined block, where Bush was associated with the positive category, and Gore was associated with the negative category (or vice versa, depending on the counter-balancing procedure), and Block 4 was the critical version of this combined block. Block 5 was a practice block consisting of Gore and Bush categorization only. Blocks 6 was a practice combined block, and was identical to blocks 3, but the politician that was associated with the positive category was now associated with the negative (and the politician that was associated with the negative category was now associated with positive), and block 7 was the critical version of this combined block. Participants completed either the traditional version of this IAT, where the labels "pleasant" and "unpleasant"

16b

10b

12b

17b

18b

26b

31b

32b

33b

29b

19-25b

were used, or the personal version, where the labels "I like" and "I don't like" were used. Error feedback was also omitted from the personalized IAT.

Participants then completed several explicit measures of their attitudes toward Bush, Gore, and other politicians. First, participants rated both Gore and Bush using several semantic differential items (Unattractive/Attractive, Bad/Good, Unpleasant/Pleasant, Foolish/Wise, and Awful/Nice). Next, participants responded to the following 5 direct comparison questions on a 7-point scale anchored by "Bush" on one endpoint and "Gore" on the other: "Who do you think is more intelligent?", "Who is more qualified to be president?", "Who do you think is more likeable?", "Whose character makes him better suited for the presidency?", and "If an election involving Bush and Gore as candidates for president were held today, for whom would you vote?" Measures of liking of Bush, Gore, and 6 filler politicians were then administered using a 7-point scale (where 0 = "Not at all" and 6 = "Very much"). Next, a feeling thermometer that included Bush, Gore, and 6 filler politicians that was analogous to Experiment 3 was administered. Participants then reported whether they voted in the 2000 presidential election, and if so, for whom they voted. Finally, participants reported their party affiliation on a 7-point scale anchored on one end by "Definitely Republican" and the other end by "Definitely Democrat," with the scale midpoint indicating "Neither/ No preference." They were then debriefed, paid, and dismissed.

#### Results and Discussion

Effects of IAT version. After dropping the first 2 trials of each block and log-transforming latencies, Gore/- and Gore/+ block means were submitted to a 2 (Block Type: Gore/+ v. Gore/-) X 2 (IAT Type: Traditional v. Personalized) ANOVA, with repeated measures on the first factor. No effects were revealed. Thus, participants were no faster on either the