American Economic Review: Papers & Proceedings 2015, 105(5): 340–345
http://dx.doi.org/10.1257/aer.p20151075

Understanding Ethnic Identity in Africa: Evidence from the Implicit Association Test (IAT)[†]

By Sara Lowes, Nathan Nunn, James A. Robinson, and Jonathan Weigel*

A fundamental feature of people's psychology is their identity—one's sense of self. A great deal of evidence suggests that individuals go through a process of choosing an identity which then heavily influences the way they behave, which social norms they adopt, and how they treat others (Akerlof and Kranton 2000). In the African context, the literature has focused on ethnicity as a central dimension of identity. While the early empirical literature worked at the macro level, more recent scholarship has moved to the individual level to better understand ethnicity, identity, and its potential economic effects. For example, Habyarimana et al. (2007, 2009) undertake an extensive set of behavioral laboratory studies with participants in Kampala, Uganda. Interestingly, they find no evidence of a differential ability to cooperate between coethnics and noncoethnics in a puzzle game that requires communication and coordination between pairs of participants. They find no evidence of differences in preferences for public goods or for differential altruism toward coethnics relative to non-coethnics using behavior in an anonymous dictator game. It is only when the identity of the dictator is observable that there is coethnic bias in giving. These findings suggest that there may not be an innate or implicit bias toward other ethnic groups, but that the social setting (e.g., social pressures or social sanctions) causes individuals to behave differently toward coethnics when observed.

Despite the contention over the importance of implicit ethnic bias, economists have yet to directly measure implicit ethnic bias. While Habyarimana et al. (2007, 2009) infer this from the outcomes of behavioral games, in this paper we argue that recent work in psychology allows us to develop a direct measure of ethnic bias or the intrinsic preference for one's own ethnic group. Our measure is a variant of the implicit association test (IAT), which is a computer-based sorting task that aims to measure individuals' implicit attitudes toward specific targets. Though the IAT has been heavily used to measure implicit attitudes toward race (black versus white) and toward gender (male versus female), it has not been widely used to examine ethnic preferences.

We implement a variant of the IAT—the single target IAT (ST-IAT)—to test for participants' implicit attitudes toward various ethnic groups. We use a participant population from Kananga, a city located in the central Democratic Republic of Congo, and measure participants' attitudes toward four ethnic groups: Luluwa, Luba, Lele, and Kuba. The sample population is the same as in Lowes et al. (2015b).

We find that the ST-IAT measures reveal evidence of an implicit own-ethnicity bias. Individuals have implicit views of their own ethnic group that are more positive than their implicit views of other ethnic groups. Using survey-based measures, we also find evidence of own-ethnicity bias in self-reported attitudes toward other ethnic groups. Interestingly, the implicit own-ethnicity bias measured by the IAT is much smaller than the explicit bias measured using survey questions.

*Lowes: Department of Economics, Harvard University, 1805 Cambridge Street, Cambridge, MA 02138 (e-mail: slowes@fas.harvard.edu); Nunn: Department of Economics, Harvard University, 1805 Cambridge Street, Cambridge, MA 02138 (e-mail: nnunn@fas.harvard.edu); Robinson: Department of Government, Harvard University, 1737 Cambridge Street, Cambridge, MA 02138 (e-mail: jrobinson@gov.harvard.edu); Weigel: Department of Government, Harvard University, 1737 Cambridge Street, Cambridge, MA 02138 (e-mail: jweigel@fas.harvard.edu).
We thank Matthew Summers for excellent research assistance throughout the project. We also thank Jonathan Yantzi for his assistance in the field and Sriram Natarajan for his guidance with the IAT programing.

[†]Go to http://dx.doi.org/10.1257/aer.p20151075 to visit the article page for additional materials and author disclosure statement(s).

We find that one shortcoming of the survey-based data is that they reveal strong enumerator-ethnicity effects. All else equal, participants report more positive attitudes toward the ethnicity of the enumerator. However, we find that this bias does not exist in the IAT measures. This is likely because the implicit attitudes elicited using the IATs are not observable to the participant, let alone the enumerator. This suggests an important potential benefit of the IAT over survey-based measures.

I. Overview of the Single-Target Ethnicity Implicit Association Test

The IAT was developed to measure an individual's implicit association between pairs of objects. To see the intuition behind the IAT, consider the following task described in Banaji and Greenwald (2013, pp. 33–34). You have to sort a deck of shuffled cards into two piles (one on your left and one on your right) based on the suit of each card. In the first task, you are asked to put hearts and diamonds in the left pile and spades and clubs in the right pile. Doing this is quite natural because hearts and diamonds are both red, and spades and clubs are both black. There is a strong association between these suits that makes sorting quick and intuitive. However, if you were asked to sort hearts and spades into the left pile and diamonds and clubs into the right, this would be much less intuitive and would take longer.

This example illustrates the basic logic of the IAT. In a standard IAT, four different types of images (like the suits) appear on a computer screen. The participant is asked to sort these images into two groups (like the two piles), one on the left side and one on the right side. If there is an underlying association between certain types of images, then some groupings will be easier to sort than others (as in the card example above). In the standard Black-White IAT, individuals observe: images of Caucasians, images of African Americans, images of good words (e.g., happy, wonderful), and images of bad words (e.g., terrible, horrible). If, for example, one has a negative implicit view of African Americans, then sorting the images of African Americans and the bad words to the same side of the screen will be easier and quicker than sorting images of African Americans and of good words to the same side of the screen. If there is no underlying association, then sorting African American and good images together should take the same amount of time as sorting African American and bad images together.

One shortcoming of the standard IAT is that one is only able to make statements about one's view of a target relative to the other target. For example, in the race IAT, one is only able to observe whether the association of good words is stronger with African American images or Caucasian images (relative to the other). One is not able to make an absolute statement about how positively the participant views African Americans or Caucasians. In addition, the standard IAT only lends itself to opposing pairs of targets e.g., black-white, male-female, etc. Many objects of interest, including ethnicity, are not naturally represented in pairs.

These shortcomings have led researchers to develop extensions of the standard IAT. For example, Bluemke and Friese (2008) have developed the single-target IAT (ST-IAT) which we use in our research. In each block of the IAT, three objects are sorted: good words, bad words, and words associated with a target (e.g., a political party, a brand of soda pop, an ethnicity, etc.). Following the same logic as the standard IAT, if the participant has a positive view of the target, then sorting will be faster when the participant has to sort the target and good words to the same side of the screen than when the participant has to sort the target and bad words to the same side of the screen.

In the summer of 2014 in Kananga, DRC, we implemented a version of the ST-IAT using ethnic groups as targets. Due to time constraints and limited participant attention span, we limited the study to four ethnicities prevalent in Kananga: Luluwa, Luba, Lele, and Kuba. Our sample includes 536 individuals. Of these, 193 are Luluwa (the most prevalent group in Kananga), 33 are Luba, 38 are Lele, and 72 are Kuba.

In this setting, a significant proportion of the population is illiterate. Therefore, we implemented a variant of the ST-IAT in which participants sorted sounds rather than written words. The ST-IAT we developed includes nine blocks in total. The first block is a practice block in which participants played a simplified version of the game, sorting only good and bad words. In every block of the ST-IAT, participants needed to obtain a 75 percent success rate in order to continue to the next block. If they did not meet this threshold, they repeated the block.

Figure 1. Single-Target IAT Screen Shot

The remaining eight blocks of the ST-IAT comprise four pairs of blocks, one pair for each of the four ethnic groups. In one block of each pair, words related to the ethnic group are sorted left and in the other, words related to the ethnic group are sorted right. In all blocks, good words are sorted left and bad words are sorted right. Each block comprises 24 trials: 8 trials with words associated with 1 of the 4 ethnic groups, 8 with good words, and 8 with bad words.[1]

Participants used a ten-inch Samsung Galaxy Tab 3 touchscreen tablet connected to headphones. A view of the screen is provided in Figure 1. In the block shown, respondents sort good words to the left, words related to the Luba ethnic group to the left, and bad words to the right. The sorting was done by pressing the red buttons on the lower left or right of the screen. Additional details of the IATs and their implementation are provided in a separate online Appendix (Lowes et al. 2015a).

The IAT measure of interest is the *D-score*, which we construct as follows (e.g., Greenwald, Nosek, and Banaji 2003; Lane et al. 2007). We ignore data from practice blocks and from any blocks that were repeated because the participant did not have an accuracy rate above 75 percent. We winsorize (i.e., truncate) the recorded latency (i.e., response time) to 3,000 milliseconds and account for incorrect responses by replacing their latency with the block mean latency plus the block standard deviation latency. The D-score measuring the positivity of the implicit association of the target is calculated as:

$$D\text{-}score = [Mean(latency^{-ve}) - Mean(latency^{+ve})]/SD(latency^{both}).$$

$Mean(latency^{-ve})$ is the recorded average response time for the block in which the ethnic group is paired with bad words, $Mean(latency^{+ve})$ is the average response time for the block in which the ethnic group is paired with good words, and $SD(latency^{both})$ is the standard deviation of the response time during both blocks. Note that if the participant is able to sort the various objects more rapidly when the ethnic group is matched with good words, then $Mean(latency^{-ve}) > Mean(latency^{+ve})$ and the *D-score* is positive. Thus, the *D-score* is increasing in the participant's implicit bias in favor of a given ethnic group.

II. Empirical Results

We begin by first testing the validity of the ST-IAT in our setting. To do this, we developed a ST-IAT for which we had strong priors about what associations we should observe. Our chosen targets of interest were food, spiders, and snakes. From initial focus groups, we confirmed that individuals liked food (not surprising) and that they disliked spiders and snakes. We also learned that they disliked snakes much more than spiders since many snakes in the area are poisonous while the spiders are not.

Results from the ST food-spiders-snakes IAT are summarized in Figure 2. Our findings confirm that within our sample of 536 individuals, the average implicit association of food is positive (and statistically different from zero), and of spiders and snakes is negative, with the association with snakes being more strongly negative than for spiders. These findings confirm that the single-target IAT succeeds in capturing participants' implicit attitudes in our setting.

Moving to the ST ethnicity IAT, our primary interests are (i) whether we observe an own-ethnicity bias in the IAT D-score, and (ii) how this compares to the explicit own-ethnicity bias found in survey-based measures.

[1] For the eight trials with ethnicity words, we randomly drew from a pool of seven ethnic words (e.g., for the Luluwa and in English): Luluwa, the Luluwa, Luluwa person, Chief of the Luluwa, Luluwa culture, Luluwa tradition, Luluwa custom.

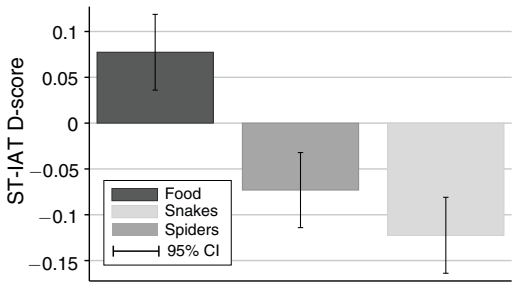

FIGURE 2. RESULTS FROM THE SINGLE-TARGET FOOD-SPIDERS-SNAKES IAT

The first set of survey questions we use asks individuals to report how close they feel to people from each of the four ethnic groups. Participants chose responses using an integer scale ranging from 0 (furthest) to 5 (closest). The second set of measures is based on respondents' reported perceptions of each ethnicity group, with the possible responses being: very negative (0), somewhat negative (1), neutral (2), somewhat positive (3), and very positive (4).

To test for the presence of a bias toward one's own ethnicity, we estimate the following equation:

$$(1) \quad y_{ie} = \alpha_i + \alpha_e + \beta I_{ie}^{OwnEthnicity} + \varepsilon_{ie}.$$

The unit of observation is an individual i's view of ethnicity e, either self-reported or measured with the ST ethnicity IAT. The variable $I_{ie}^{OwnEthnicity}$ is an indicator variable that equals 1 if the ethnicity of the target ethnic group e is the same as the (self-reported) ethnicity of individual i. The equation also includes fixed effects for the four ethnic groups, α_e, as well as individual fixed effects, α_i. The baseline sample includes 536 individuals belonging to 22 different ethnic groups. All of the results reported are very similar if we restrict the sample to the 336 participants that belong to one of the four target ethnic groups.

Estimates of equation (1) are reported in columns 1–3 of Table 1. In columns 1 and 2, the dependent variable is participants' self-reported views of each of the four ethnic groups, and in column 3, it is the IAT D-score measure. For all three outcomes one observes a statistically significant own-ethnicity bias. Interestingly, the own-ethnicity bias is much stronger for the survey questions than for the IAT score. While the size of the own-ethnicity bias when using the two survey measures is 0.74 and 0.58 standard deviations (columns 1 and 2), the size of the effect when the D-score is used is 0.14 standard deviations (column 3). This suggests that individuals' implicit ethnic bias toward their own group is less strong than their explicit self-reported bias.

This result dovetails nicely with Habyarimana et al. (2007, 2009), who find evidence of an ethnic bias among participants only when their actions are observed. Their findings suggest the existence of an explicit bias but no implicit bias. Here, we also find evidence for an explicit own-ethnicity bias that is larger than the implicit bias. However, in contrast to Habyarimana et al. (2007, 2009), we find that the implicit bias is not zero. Our IAT-based measure is able to identify the existence of an implicit bias toward one's own ethnic group.

There are a number of potential concerns with survey-based questions that ask respondents about their views of various ethnic groups. One, in particular, is that individuals' self-reported views may be influenced by the ethnicity of the enumerator. For example, it may be more difficult to report that one does not view the Luluwa positively when the enumerator is Luluwa.

One potential benefit of the IAT over survey-based measures (or experiment-based measures) is that the enumerator does not observe the participant's performance in the IAT. Further, since the IAT score is based on split-second response differences, it is highly unlikely that the enumerator could detect an ethnic bias even if the enumerator were watching.

To assess the extent to which enumerator ethnicity affects our outcomes of interest, we include in equation (1) an indicator variable that equals 1 if the ethnicity of the enumerator is the same as the target ethnicity e. The estimates, with the two survey-based measures as outcomes are reported in columns 4 and 5 of Table 1. They show that the ethnicity of the enumerator does matter. Participants' self-reported views of an ethnic group are more positive when the enumerator belongs to that ethnic group. In contrast, the enumerator's ethnicity does not affect the ST ethnicity IAT D-score. As reported in column 6, when the D-score is the dependent variable, the coefficient for the enumerator

Table 1—Baseline Results

	Dependent variable					
	Feeling of closeness, 0–5 (1)	Positivity of perception, 0–4 (2)	ST ethnicity IAT D-score (3)	Feeling of closeness, 0–5 (4)	Positivity of perception, 0–4 (5)	ST ethnicity IAT D-score (6)
Same ethnicity as participant	0.942***	0.597***	0.062**	0.951***	0.606***	0.061**
	(0.068)	(0.059)	(0.029)	(0.068)	(0.059)	(0.029)
Same ethnicity as enumerator				0.205**	0.212***	−0.011
				(0.094)	(0.082)	(0.040)
Mean of the dependent variable	3.19	2.90	0.035	3.19	2.90	0.035
SD of the dependent variable	1.26	1.04	0.44	1.26	1.04	0.44
Individual fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Ethnic group fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,133	2,133	2,133	2,133	2,133	2,133
R^2	0.52	0.47	0.29	0.53	0.47	0.29

Notes: The unit of observation is the view of an individual (536 individuals in total) toward an ethnic group (four ethnic groups in total: Luluwa, Luba, Lele, and Kuba). The dependent variable in columns 1 and 4 is based on participants' reports of how close they feel to those from each of the four ethnic groups. They are asked to choose answers ranging on an integer scale from 0 (furthest) to 5 (closest). The dependent variable in columns 2 and 5 is based on a survey question that asks respondents to report their perceptions of each ethnicity group. The possible answers were: very negative (0), somewhat negative (1), neutral (2), somewhat positive (3), and very positive (4). The dependent variable in columns 3 and 6 are the D-score from the ST ethnicity IAT. Coefficients are reported with standard errors in parentheses.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

ethnicity indicator variable is very close to zero and statistically insignificant. These results suggest that a potential benefit of using the IAT to measure attitudes toward ethnic groups and to detect ethnic bias is that it is less susceptible to influence arising from the observability of participant responses.[2]

III. Conclusion

We have examined a new measure of attitudes toward ethnic groups and have found evidence that the single-target IAT is a valid measure of ethnic views. We use the measure to test for an own-ethnicity bias in our sample and find that the bias we estimate using the single-target IAT is smaller in magnitude and statistical strength relative to the bias found when using self-reported views toward ethnic groups from surveys. We also identify an important benefit of using IATs to measure ethnic bias. With the survey questions, the ethnicity of the enumerator has a strong effect on participant responses. In contrast, the IAT appears immune to such manipulation, perhaps due to the unobservability of sorting latencies. These findings suggest that the IAT is a useful complementary tool in studying ethnicity in Africa and elsewhere.

[2] We also find evidence that the order of activities and surveys, which we randomly assigned, affects explicit ethnic bias measured using the survey questions, but does not affect implicit ethnic bias measured using the IAT. The estimated coefficient of interest, β, in equation (1), is sensitive to the order of activities when survey measures are used but not when the IAT D-score is used.

REFERENCES

Akerlof, George A., and Rachel E. Kranton. 2000. "Economics and Identity." *Quarterly Journal of Economics* 115 (3): 715–53.

Banaji, Mahzarin R., and Anthony G. Greenwald. 2013. *Blindspot: Hidden Biases of Good People*. New York: Delacorte Press.

Bluemke, Matthias, and Malte Friese. 2008. "Reliability and Validity of the Single-Target

IAT (ST-IAT): Assessing Automatic Affect Towards Multiple Attitude Objects." *European Journal of Social Psychology* 38 (6): 977–97.

Greenwald, Anthony G., Brian A. Nosek, and Mahzarin R. Banaji. 2003. "Understanding and Using the Implicit Association Test: I. An Improved Scoring Algorithm." *Journal of Personality and Social Psychology* 85 (2): 197–216.

Habyarimana, James, Macartan Humphreys, Daniel N. Posner, and Jeremy M. Weinstein. 2007. "Why Does Ethnic Diversity Undermine Public Goods Provision?" *American Political Science Review* 101 (4): 709–25.

Habyarimana, James, Macartan Humphreys, Daniel N. Posner, and Jeremy M. Weinstein. 2009. *Coethnicity: Diversity and the Dilemmas of Collective Action*. New York: Russell Sage.

Lane, Kristin A., Mahzarin R. Banaji, Brian A. Nosek, and Anthony G. Greenwald. 2007. "Understanding and Using the Implicit Association Test: IV. What We Know (So Far) about the Method," In *Implicit Measures of Attitudes,* edited by Bernd Wittenbrink and Norbert Schwarz, 59–102. New York: The Guilford Press.

Lowes, Sara, Nathan Nunn, James A. Robinson, and Jonathan Weigel. 2015a. "Appendix for: Understanding Ethnic Identity in Africa: Evidence from the Implicit Association Test (IAT)." http://scholar.harvard.edu/files/nunn/files/iat_appendix_1.pdf.

Lowes, Sara, Nathan Nunn, James A. Robinson, and Jonathan Weigel. 2015b. "The Evolution of Culture and Institutions: Evidence from the Kuba Kingdom." Unpublished.