Online Web Appendix: Is Ethnicity Identifiable? Lessons from an Experiment in South Africa

Adam Harris

(ash338@nyu.edu)

Michael Findley

(mikefindley@austin.utexas.edu)

February 8, 2012

Details of Ethnic Identity Measure and Other Survey Questions

Questions adopted from the Multigroup Ethnic Identity Measure (Phinney 1992)

Scoring Pre-study Questionnaire

Ethnic identity is measured on a 1 (weak ethnic identity) to 4 (strong ethnic identity) (Phinney 1992, 164). First, people responded to the statements listed below based on the following scale: 1=strongly disagree, 2=somewhat disagree, 3=somewhat agree, 4=strongly agree. We calculated a score on the 1 to 4 scale, by summing and then averaging the numerical responses. However, questions 6 and 7 generate opposite numerical responses. For questions 6 and 7, an answer of "strongly agree" is the answer a weaker ethnic would give, but "strongly agree" is associated with a 4. Therefore, we first reversed the subjects' answers for questions 6 and 7 before summing and averaging (173). This means that for questions 6 and 7, if the respondent puts a 2, then the answer is changed to a 3 before summing and averaging the responses. This score is the *ethnic strength* independent variable in our regression analysis.

1. I have spent time trying to find out more about my own ethnic group, such as its history, traditions, and customs _____

2. I am active in organizations or social groups that include mostly members of my own ethnic group _____

3. I have a clear sense of my ethnic background and what it means for me _____

4. I think a lot about how my life will be affected by my ethnic group membership _____

5. I am happy that I am a member of the group I belong to _____

6. I am not very clear about the role of my ethnicity in my life _____

7. I really have not spent much time trying to learn more about the culture and history of my ethnic group _____

8. I have a strong sense of belonging to my ethnic group _____

9. I understand pretty well what my ethic group membership means to me, in terms of how to relate to my own group and other groups _____

10. In order to learn more about my ethnic background, I have often talked to other people about my ethnic group _____

11. I have a lot of pride in my ethnic group and its accomplishments _____

12. I participate in cultural practices of my own group, such as special food, music, or customs _____

13. I feel a strong attachment toward my own ethnic group _____

14. I feel good about my cultural or ethnic background _____

Questions adopted from Habyarimana et al (2007) and used in our pre-study questionnaire

Pre-study Questionnaire

Background Information: Respondent Number _____

1. Age:

2. Gender:

3 Where do you stay now (include extension or section number; e.g. Mdantsane NU2)?

4 What province were you born in?

5 How many years have you lived in Mdantsane?

6 How many years have you lived in your current house?

7 What is your marital status?

8 With which ethnic/cultural group do you belong to?

9 What is your first language?

10 What other languages can you speak?

11 What is your religion?

12 How often do you attend a place of worship (e.g. church, mosque, temple, etc)?

13 How much schooling have you done?

14 How much schooling did your father do?

15 How much schooling did you mother do?

16 Which of the following do you do most of the time?

17 How often do you see/read the following types of media (television, radio, newspaper, internet)?

18 Did you vote in the most recent parliamentary elections?

19 What is the roof of your house made of?

20 What is the main source of drinking water in your household?

21 Does your household have any of the following? (Electric iron, refrigerator, mattress or bed, radio, watch or clock, sewing machine, modern stove, bicycles, motorcycle, car or truck, electricity)

22a Hows many children do you have?

22b How many children do you have between the ages of 5 and 18?

22c How many of your kids are attending school?

Summary Statistics and Additional Results Tables

Summary Statistics

Variable	Mean	Standard Deviation	Minimum	Maximum
Guess	0.27	0.45	0	1
Ethnic Strength	3.42	0.34	2.5	4
Subject Age	30.06	11.32	16	59
Subject Gender	0.53	0.5	0	1
Schooling	5.44[+]	1.17	4	7
Schooling Father	4.47	2.07	1	8
Schooling Mother	4.5	1.51	2	8
Years in Mayibuya	20.24	11.77	1	59
Religiosity	2.61	1.6	1	7
Ethnic Representer Age	29.74	11.89	16	72
Ethnic Representer Gender	0.54	0.5	0	1
Ethnic Strength of Ethnic Representer	3.35	0.38	2.36	4
Gender Match	0.5	0.5	0	1
Age Distance	12.42	10.75	0	56
Ethnic Strength Match	0.38	0.49	0	1
Ethnic Match (Both Subject and Representer are Xhosa)	0.22	0.42	0	1

[+]Converting 5.44 to a qualitative interpretation shows that on average, the subjects have at least completed high school because 5 represents "completed high school" and a 6 represents "completed some university or college".

Percentage of Correct Guesses by Level of Ethnic Strength

Ethnic Strength	(1) All Conditions	(2) Photograph	(3) First Name	(4) Surname	(5) Greet Own Language	(6) Greet English
Bottom Third	26.0	17.7	25.1	28.8	49.0	15.5
Middle Third	27.8	18.1	31.2	23.1	46.5	17.3
Top Third	28.4	18.4	28.2	32.6	48.0	16.4

Ethnic Strength	(7) Signal Own (True) ID	(8) Signal Own w/Own Symbol	(9) Signal Own w/Other Symbol	(10) Signal Other (Fake) ID	(11) Signal Other ID w/Other Symbol	(12) Signal Other ID w/Own Symbol
Bottom Third	63.8	73.9	65.4	5.9	8.6	11.4
Middle Third	73.9	77.6	66.2	5.4	8.1	15.1
Top Third	78.7	86.2	75.0	2.7	5.8	5.1

The table shows how the different levels of ethnic strength effect successful identification under the various treatment conditions. Each cell contains the percentage of successful guesses for each ethnic strength category. As is consistent with the regression results, the top third percentile has the *highest* rate of successful identification for "All Conditions" and "Surname" This is also the case for conditions (7), (8), and (9). Also, under conditions (10) and (11) the top thrid percentile of ethnic strength has the *lowest* success rate of identification, which is also consistent with the regression results.

Control Variables from Table 2 Model

Dependent Variable: Correct Guesses

VARIABLES	All Conditions
Subject Characterisics	
Subject: Age	-0.002
	(0.006)
Subject: Gender	0.189
	(0.168)
Subject: Education	0.023
	(0.058)
Subject: Father's	-0.013
Education	(0.037)
Subject: Mother's	-0.000
Education	(0.035)
Subject: Years in	0.006
Mayibuya	(0.006)
Subject: Religious	-0.033
Activity	(0.033)
Representer Characteristics	
Representer Age	-0.007*
	(0.004)
Representer Gender	0.090
	(0.069)
Representer	0.112
Ethnic Strength	(0.095)
Dyadic Variables	
Gender Match	0.022
	(0.071)
Age Distance	0.004
	(0.004)
Ethnic Strength Match	-0.039
	(0.080)
Ethnicity Match	1.411***
	(0.114)
Order Controls	
Experimental Order	-0.004
	(0.003)
Order Squared	0.000**
	(0.000)
Constant	-3.454***
	(0.711)
Observations	7795

Complete Table 3 With All Covariates

VARIABLES	(1) Photograph	(2) First Name	(3) Surname	(4) Greet Own Language	(5) Greet English
Subject: Strength Score	0.145	0.608	0.757**	0.333	-0.031
	(0.153)	(0.453)	(0.377)	(0.422)	(0.480)
Subject Characterisics					
Subject: Age	-0.003	-0.022	-0.019	-0.013	-0.0376*
	(0.006)	(0.018)	(0.017)	(0.013)	(0.022)
Subject: Gender	0.213*	0.062	0.104	-0.316	-0.933**
	(0.119)	(0.270)	(0.403)	(0.326)	(0.382)
Subject: Education	-0.008	0.243**	-0.095	0.328***	0.095
	(0.050)	(0.117)	(0.140)	(0.116)	(0.172)
Subject: Father's Education	-0.014	0.014	-0.104	0.095	-0.009
	(0.031)	(0.072)	(0.073)	(0.065)	(0.081)
Subject: Mother's Education	-0.051*	0.029	0.035	0.075	0.156
	(0.030)	(0.087)	(0.086)	(0.092)	(0.143)
Subject: Years in Mayibuya	0.002	0.003	0.021	0.007	0.043**
	(0.005)	(0.014)	(0.015)	(0.010)	(0.019)
Subject: Religious Activity	-0.021	0.061	-0.080	-0.020	0.152
	(0.032)	(0.074)	(0.089)	(0.112)	(0.115)
Represener Characteristics					
Represener Age	-0.006	-0.031**	-0.009	-0.020*	0.015
	(0.005)	(0.015)	(0.012)	(0.010)	(0.016)
Represener Gender	0.039	0.284	-0.070	0.156	0.346
	(0.104)	(0.263)	(0.271)	(0.245)	(0.348)
Represener Ethnic Strength	0.112	0.663	0.320	0.446	-0.130
	(0.129)	(0.425)	(0.367)	(0.301)	(0.487)
Dyadic Variables					
Gender Match	0.026	0.061	0.133	-0.345	0.491
	(0.105)	(0.238)	(0.279)	(0.238)	(0.346)
Age Distance	0.004	0.016	-0.014	0.007	-0.011
	(0.006)	(0.016)	(0.012)	(0.013)	(0.022)
Ethnic Strength Match	-0.056	-0.166	0.106	0.163	0.179
	(0.100)	(0.231)	(0.288)	(0.252)	(0.357)
Ethnicity Match	1.407***	2.439***	1.744***	2.845***	1.940***
	(0.132)	(0.368)	(0.287)	(0.418)	(0.399)
Order Controls					
Experimental Order	-0.006*	-0.004	-0.010	-0.003	-0.018
	(0.003)	(0.011)	(0.009)	(0.010)	(0.016)
Order Squared	0.000*	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	-2.219***	-6.418***	-3.386*	-4.794***	-3.476
	(0.669)	(2.243)	(1.870)	(1.847)	(2.122)
Observations	3906	404	396	398	351

Complete Table 5 With All Covariates

VARIABLES	(6) Signal Own (True) ID	(7) Signal Own w/Own Symbol	(8) Signal Own w/Other Symbol	(9) Signal Other (Fake) ID	(10) Signal Other ID w/Other Symbol	(11) Signal Other ID w/Own Symbol
Subject: Strength Score	0.942	1.481*	0.952	-1.375*	-0.266	-1.129
	(0.740)	(0.798)	(0.720)	(0.729)	(0.827)	(0.877)
Subject Characterisics						
Subject: Age	0.010	0.020	0.016	-0.097	0.024	0.026
	(0.025)	(0.026)	(0.022)	(0.082)	(0.032)	(0.042)
Subject: Gender	0.657	0.283	0.805	-0.103	0.718	0.157
	(0.661)	(0.554)	(0.534)	(0.703)	(0.535)	(0.545)
Subject: Education	-0.004	-0.215	-0.066	0.320	0.120	0.209
	(0.234)	(0.211)	(0.196)	(0.421)	(0.235)	(0.223)
Subject: Father's Education	-0.143	-0.011	0.030	-0.145	0.143	0.028
	(0.125)	(0.129)	(0.116)	(0.147)	(0.145)	(0.107)
Subject: Mother's Education	0.042	0.019	0.010	-0.012	-0.106	0.159
	(0.181)	(0.195)	(0.163)	(0.237)	(0.198)	(0.205)
Subject: Years in Mayibuya	0.015	0.015	0.021	0.024	-0.046	-0.033
	(0.022)	(0.022)	(0.022)	(0.036)	(0.028)	(0.024)
Subject: Religious Activity	-0.178	-0.172	-0.173	0.042	0.087	0.303
	(0.164)	(0.162)	(0.152)	(0.148)	(0.172)	(0.193)
Representer Characteristics						
Representer Age	-0.002	0.004	0.002	-0.097*	-0.005	-0.005
	(0.013)	(0.012)	(0.013)	(0.055)	(0.015)	(0.017)
Representer Gender	-0.024	0.463**	0.314	0.280	-0.171	0.535
	(0.243)	(0.223)	(0.203)	(0.498)	(0.352)	(0.377)
Representer Ethnic Strength	0.040	-0.267	-0.234	0.334	-0.347	0.617*
	(0.217)	(0.403)	(0.294)	(0.506)	(0.507)	(0.355)
Dyadic Variables						
Gender Match	0.119	0.260	-0.062	-0.389	-0.068	0.319
	(0.224)	(0.232)	(0.194)	(0.543)	(0.378)	(0.366)
Age Distance	0.023	-0.008	0.012	0.049	-0.014	0.032
	(0.015)	(0.015)	(0.012)	(0.061)	(0.022)	(0.022)
Ethnic Strength Match	0.281	-0.032	0.019	-0.951*	-0.140	-0.663
	(0.225)	(0.287)	(0.237)	(0.570)	(0.420)	(0.404)
Ethnicity Match	0.574**	0.451	0.067	1.611**	1.099**	1.032**
	(0.267)	(0.368)	(0.283)	(0.649)	(0.490)	(0.481)
Order Controls						
Experimental Order	-0.008	0.004	0.017*	-0.019	-0.001	0.013
	(0.010)	(0.011)	(0.009)	(0.029)	(0.015)	(0.013)
Order Squared	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	-2.577	-2.811	-3.442	4.123	-1.600	-4.944*
	(2.652)	(3.363)	(2.698)	(3.394)	(3.110)	(2.736)
Observations	403	374	412	379	403	369

Pooling Conditions to Consider Joint Effects

We also considered a number of models in which we pooled various conditions. The results are reported below. When pooling all of the sign conditions, the result is positive and significant, indicating that stronger ethnics are generally better at identifying others.

We then pooled the signal conditions in five different ways:

1. All signal conditions
2. Signal conditions without any symbol (confirming or disconfirming)
3. Signal conditions with any symbol
4. Signal conditions with supportive symbols
5. Signal conditions with contradictory symbols

These results suggest that the signal conditions generally aid accurate identification, especially when supportive symbols are displayed. When only contradictory symbols are included, guesses appear to be less accurate.

Ethnic Strength coefficient estimates based on pooling, along with statistical significance:

1. All sign conditions: 0.266*
2. All signal conditions: 0.448*
3. Signal conditions without symbols: 0.366
4. Signal conditions with symbols: 0.473*
5. Signal conditions with supportive symbols: 0.608**
6. Signal conditions with contradictory symbols: 0.341