Ethnic Measurement and Identification

TOM W. SMITH

National Opinion Research Center

Like religion and political party preference ethnicity is not a simple all or nothing proposition.[1] Researchers have long recognized that a person's level or intensity of identification with a particular ethnicity can vary from a weak-nominal association to a strong-committed association. As Abner Cohen observed, "Again we must remember that ethnicity is a matter of degree. There is ethnicity and ethnicity. . . . Unless we recognize differences in degree of manifestation we shall fail to make much progress in the analysis of ethnicity. To put it in the idiom of research, ethnicity is a variable" (Cohen, 1974, p. xiv). To measure the strength of the ethnic affiliation various identification scales have been used. Some are keyed to a particular ethnic group (e.g., Uyeki, 1960; Segalman, 1967; Masuda, Matsumoto, & Meredith, 1970) while others attempt to serve as transethnic measures (Pavlak, 1976; Driedger, 1975; Plax, 1972). The identification measures also differ considerably in the number of dimensions they include and the number of items measuring each dimension. Plax, for example, has a measure of passive identification ("We talked a few minutes ago about your nationality background. Do you ever think of yourself as being [NAME OF GROUP]?") and a measure of saliency ("How important was being [NAME OF GROUP] in your family? Is it important to you?"). Dreidger on the other hand taps six dimensions: religion, endogamy, friendships, language use, parochial education, organizational membership, and media use with four items each that measure both attitudes and behavior. The collective drawback of all the ethnic identification studies is that they all start with individuals a priori defined as members of particular ethnic groups and then ask the question, how strong is this person's identity with his ethnicity? This ignores the important prior questions as to how a person knew or decided he was a member of a particular ethnic group in the first place, and how a researcher determined that he was a member of a particular ethnic group. The first step in studying ethnic identification is determining what (if any) ethnicity a person identifies with rather than whether he identifies with it strongly or weakly. To consider this question we have to explore how ethnicity is measured.

Basically three approaches to ethnic measurement can be delineated, the natal, the behavioral, and the subjective. The natal approach attempts to identify a person's ethnicity by inquiring about a person's place of

birth, the places of birth of his parents, grandparents, and so on. The behavioral approach defines a person's ethnicity according to some practice or affiliation of the person such as language spoken or membership in certain voluntary groups. The subjective approach simply asks the person what ethnicity he considers himself or where his ancestors came from.[2]

The nativity approach to ethnicity is typified by the traditional item used by the Bureau of the Census. It asked the place of birth of the respondent, his mother, and his father (see below). A variant of

13a. WHERE WAS THIS PERSON BORN? If born in hospital, give State or country where mother lived. If born outside U.S., see instruction sheet: distinguish Northern Ireland from Ireland (Eire).
☐This State
   OR _____
         (Name of State or foreign country; or Puerto Rico, (Guam, etc.)

14. WHAT COUNTRY WAS HIS FATHER BORN IN?
☐United States
   OR _____
         (Name of foreign country; or Puerto Rico, Guam, etc.)

15. WHAT COUNTRY WAS HIS MOTHER BORN IN?
☐United States
   OR _____
         (Name of foreign country; or Puerto Rico, Guam, etc.)

this approach, used by the Michigan Election studies, asks parallel information about the respondent and his parents, but then inquires about the general ancestral origins of respondents who are third generation or above (see below). This begins to shift from being an objective measuring of nativity to a subjective or self-identification measure.

1. Where were you born? (IF UNITED STATES) Which state?
2. Were both your parents born in this country?
   —If response to Q. 2 was "no":
   2A. Which country was your father born in?
   —If response to Q. 2 was "no":
   2B. Which country was your mother born in?
   —If response to Q. 2 was "yes" or "Don't know":
   2C. Do you remember which country your family came from originally on your father's side?
   —If response to Q. 2 was "yes" or "don't know":
   2D. Do you remember which country your family came from originally on your mother's side?

Inquiring about "which country your family came from" does not mea-

sure the national origins of the respondent's ancestors as a whole, but rather elicits a single origin from among potentially several different ancestral lines. A strict (although obviously impractical) natal approach would inquire about the place of birth of all ancestors until all lines were traced back to the country of origin. Another mixed approach (used by National Opinion Research Center (NORC), Survey Research Service (SRS 857) asks "Please tell me what country most of your ancestors came from?" PROBE FOR ONE COUNTRY MOST CAME FROM. "What country did *most* of person's (your father/your father's father/your mother's father) ancestors come from?" PROBE FOR *ONE* COUNTRY *MOST* CAME FROM. While it sets an objective standard by which the country of origin is to be selected (plurality origin), it actually does not collect any information to determine if this criterion is being employed. This shift from the objectivity of place of birth to subjective identification proceeds one step further in a standard Michigan Election study question which asks, "In addition to being an American, what do you consider your *main* ethnic or nationality group?" This question emphasizes one's "main" background but does not make clear how this definition is to be operationalized. It further moves away from the place of birth definition by referring to "ethnic or nationality group" rather than to country of origin in a geopolitical sense. Similar in kind is the Current Population Survey item that asks, "What is _____ 's origin or descent?" It establishes no criteria and uses the somewhat less specific terms "origin and descent." Even more clearly subjective is the General Social Survey question which asks, "From what countries or part of the world did your ancestors come? *IF MORE THAN ONE COUNTRY IS NAMED:* Which of these countries do you feel closer to?" Not only is no criterion specified to choose between origins in the initial question, but in the follow-up question people giving multiple origins are told to use a subjective standard, feeling closer to, rather than choosing a particular lineage or their most frequent origin. Another example of a subjectively directed ethnicity question is the following from a study of ethnic voting in Buffalo (Plax, 1972), "Since we're talking about your background, let me ask you your feelings about it. Sometimes our nationality backgrounds make us think of ourselves not only as Americans, but as related to other countries, and we call ourselves French, or English, or Swiss. Thinking of your background, what would you call yourself?" As in prior examples, the respondent is asked to choose a national origin without any precise reference to actual places of births of ancestors and the emphasis is on subjective identification, "your feelings," "what would you call yourself" (see also Abramson, 1975).

Clearly from these above examples the distinction between the nativity and subjective approaches is often a fine one. In general one passes from

the nativity approach when one moves from asking information about the place of birth of specific persons (the respondent and his ancestors) to nonspecific information about one's background, descent, ethnicity, nationality, or origin.

More clearly separated from the nativity or subjective approaches is the behavioral approach which classifies a person according to some practice or affiliation such as language spoken or membership in certain voluntary associations. An example of the language approach comes from the 1970 federal census which inquires, "What language, other than English, was spoken in this person's home when he was a child?" The affiliation approach is commonly used when a list sample is used to select respondents.[3] Under this method, membership in the association that the list represents becomes the definition of nationality. This might include congregations, mutual benefit societies, or other groups (for examples, see Vrga, 1971; Masuda et al., 1970; Barton, 1975). Also included in this approach are lists based on such documents as baptismal and marriage registers (which frequently include persons not actually members of such congregations). Another hypothetical example would be a survey that asked a series of ethnic orientation questions (such as foods eaten, music preferred, etc.) and then assigned ethnicity according to the responses.

All of these approaches have strengths and weaknesses. The nativity approach calls for the collection of precise data on national origins and when multiorigins are reported gives the researcher the option of classifying the person according to various possible criteria (e.g., father's side, mother's side, majority of ancestors, etc.). The chief problem[4] is that for most of the population such information is impossible to collect. Approximately 56% of the adult population report that all four of their grandparents were native born. For this large segment of the population the respondent would have to be able to identify the place of birth of at least eight great-grandparents (and of many more if all eight great-grandparents were not immigrants). This is beyond the knowledge of most people.[5] The behavioral approach also suffers from being restrictive. The 1970 Census reported that only 12 percent did not report English as their "mother tongue." The affiliation approach is even more restrictive with only 3% of adults reporting membership in "nationality groups" (Davis, Smith, & Stephenson, 1978, p. 115). In brief, it appears that neither the nativity nor the behavioral approach can be used to determine the national origins of a large number of Americans. That leaves the subjective approach. As we will see however, while this approach does maximize ethnic identification, even it fails to identify people in a large proportion of cases.

To examine the subjective approach, the questions used by the Current Population Survey (CPS), the Michigan Election studies, and the NORC General Social Surveys will be considered. The CPS item asks for the

person's "origin or descent" and supplies a list of 15 categories to choose from: five Hispanic categories, seven European groups, one racial group, "Negro," and two residual groups, "Other" and "Don't know" (see below).[6]

What is . . .'s origin or descent?
(Show Flash Card or read list)

German	0	Mexican, Chicano	0
Italian	0	Puerto Rican	0
Irish	0	Cuban	0
French	0	Central or So. Amer.	0
Polish	0	Other Spanish	0
Russian	0	Negro	0
English, Scot., Welsh	0	Other (SPECIFY)	0
		Don't know	0

Except in a few instances respondents giving multiple backgrounds are coded as "Other." Respondents who reply that they are "American" are told "that ethnic background is determined by the lineage of a person's ancestors" and probed for a national origin. Only if the respondent continues to assert that he is "American" is this accepted as a specified "Other" response. Similarly, religious affiliations are not accepted as national origins and an appropriate response is probed for. If the respondent persists in using a religious affiliation it is also recorded as a specified "Other." Looking at the March 1972 CPS as an example we find the following distribution of national origins:

Other	.312
English, etc.	.144
German	.125
Negro	.103
Don't know	.086
Irish	.080
Spanish	.045
Italian	.043
French	.026
Polish	.025
Russian	.011

1.000
(204,840)

One obvious problem with this item is that the largest group is the residual category "Other," which include (1) those with unlisted nationalities, (2) those with multiple nationalities, (3) those with "American" nationality,

and (4) those with a religious affiliation. Since these various types of "Other" nationality are not coded separately, it becomes difficult to work with a variable on which nationality is unknown or unspecified for almost 40% of the population.

Another subjective item is that used since 1972 by the Michigan Election studies. As we mentioned earlier, it actually consists of three different types of questions. Nativity questions about the respondent and his parents, a general origin question about his paternal and maternal lineages, and a subjective question on ethnicity/nationality. Looking at the subjective question first, we find that a large and apparently growing proportion of the non-black population is unable to give a clear national origin (see Table 1). Approximately 27% in 1972, 30% in 1974, and 36% in 1976 were unable to give any foreign national origin. Even among those with some foreign nationality in mind not all could give a specific origin. Approximately 8 to 12% gave a multinational region as their origin (such as Scandinavia, Central Europe, or "Asia"). Another 1 to 3% gave a combination of backgrounds that were not codeable even in broad regional groupings. As a result only 61% in 1972, 57% in 1974, and 54% in 1976 gave a clear national origin.

Even when the battery of questions on nativity are considered along

TABLE 1
ETHNIC IDENTIFICATION ON THE MICHIGAN ELECTION SUBJECTIVE ITEM (NON-BLACKS ONLY)[a]

	Year		
Identification	1972	1974	1976[b]
None	.078[c]	.096	.196
Don't know	.091	.100	.118
American	.100	.100	.050
One nationality	.608	.567	.536
Group nationality	.094	.118	.084
Multiple nationality	.029	.019	.016
Total	1.000	1.000	1.000
N	2397	1404	1954

[a] Blacks were not asked the ethnic question in 1976.
[b] It would appear that codes "none" and "American" were handled differently in 1976 than in 1972 or 1974.
[c] The code "none" also includes missing values. Based on the average number of missing values on the five or six questions of similar type and format that came before or after the subjective ethnicity in the interview, the average number of missing cases was estimated as 19 in 1972, 36 in 1974, and 17 in 1976. This reduces the total number of cases to 2378, 1368, and 1937, respectively. Removing these from the "nones" reduces their share to .071, .072, and .189, respectively, and increases the share of the other codes slightly.

with the nativity question (place of birth of respondent or family), a size-able number of nonresponses remains (11 to 14%) (see Table 2). The reduction is great however, from 26.9 to 10.7% in 1972, 29.6 to 14.2% in 1974, and from 36.4 to 11.2% in 1976. This indicates that many people unable or unwilling to name a main nationality actually know some con-crete information about their national origins. The combined ethnic iden-tity scale also confirms, however, that many people do not have a single or dominant identification and that this may be increasing (single versus mixed in 1972 = .698/.195, in 1974 = .640/.219, and in 1976 = .653/.236). In sum, it appears that by using the Michigan combination of subjective and natal items some type of ethnic identity can be determined for about 86 to 89% of the population although this is achieved in part by the use of regional or even continental categories.

On the General Social Surveys (GSS) subjective question the person is asked "From what countries or part of the world did your ancestors come?" and if giving multiorigins is asked "Which one of these countries do you feel closer to?" A person falls out of the question if he either (1) can not give any origin to the original query, (2) replies "American," "United States," or some other similar reference, or (3) can not choose between countries in the second query. In Table 3 the data show that 53% have one clear identification and another 25% are aware of multiple backgrounds but can select a dominant ethnicity. This gives a total of 78% with a national background coded. Of the remaining 22%, 12% are aware

TABLE 2

ETHNIC IDENTIFICATION ON THE MICHIGAN ELECTION SUBJECTIVE
AND NATIVITY ITEMS COMBINED (NON-BLACKS ONLY)[a]

Identification	Year		
	1972	1974	1976
None chosen	.107	.142	.112
One nationality	.698	.640	.653
Two or more nationalities	.093	.093	.140
Group nationality	.095	.123	.093
Multiple nationality	.007	.003	.003
Total	1.000	1.001	1.001
N	2397	1404	1954

[a] People with "None chosen" failed to give any non-American nationality on the sub-jective question or any national origin for themselves or their relatives on the nativity questions. People with one nationality gave a single origin on the subjective question or the nativity questions. People with two or more nationalities gave more than one country of origin for their ancestors on the nativity questions and gave no ethnicity on the subjective question. People with group and multiple nationality gave only these collective identifica-tions.

TABLE 3

ETHNIC IDENTIFICATION OF THE GSS SUBJECTIVE ITEM, 1972–1977 POOLED (WHITES)

Identification	
Can't name any country	.086
American	.015
Names more than one country, cannot choose	.117
Names more than one country, chooses	.253
Names one country	.528
N	7983

of national origins, but can not choose between them and the remaining 10% can not name any country or name America as their origin.

In terms of maximizing the number of identifications this question with a 78% take rate works better than the Michigan subjective item with take rates falling from 73% in 1972 to 64% in 1976, but not as well as the combined Michigan questions (average take rate of about 88%). On the other hand, since this question uses individual country identifications (with some exceptions) rather than national plus regional groupings, it is more precise in its identifications. Thus while the GSS codes a particular country for 78% of the population, the Michigan subjective question does so for only 54 to 61% and the combined Michigan questions for 65 to 70%. Finally, this question provides some measure of ethnic homogeneity. It records who gives more than one origin and whether a dominant origin can be chosen from the given origins, but does not record other than the chosen identification. The subjective Michigan item, on the other hand, does not have a clear measure of homogeneity (the multiple nationality codes do indicate multiple origins, but the group codes indicate either that more than one country was mentioned within the region (multiorigins) or that reference was made to the larger unit only (vague origins). The natal origin questions potentially have a lot of valuable information on homogeneity. If a person was born in the United States, it asks about the place of birth of both parents. If both parents were native born, then inquiry was made about the origin of the maternal and paternal lines. As a result, mixed ethnic background can be measured if parents came from different countries or if paternal and maternal lines had different origins. In addition, unlike the GSS question, it is possible to tell more precisely where the mixture occurs and its degree and it is possible to tell what countries are involved. Finally, the nativity questions on the Michigan elections surveys also collect information on immigration generation.[7] This additional type of information is not available from the GSS item, but starting with the 1977 survey a separate immigration question was also asked. It measures immigration generation, but does not provide addi-

tional direct information on ethnic identification or homogeneity. In sum, it appears that the subjective approaches can determine the ethnicity of a substantially higher proportion of people than the strict natal or behavioral approaches. Even this type of approach, however, fails to provide an identification for a large number of people.

To determine why it is that many people cannot give an ethnic origin, analysis was carried out on the General Social Surveys. The dependent variable, ethnic identification, was collapsed into three categories: (1) those naming no country plus Americans, (2) those naming or choosing one country, and (3) those unable to choose between countries. In a sense, these can be thought of as people with too little ethnicity, the "right amount" of ethnicity, and too much ethnicity.

Initially this ethnic identification variable was related to four dimensions—familial and life cycle, socioeconomic level, heritage, and generation. The familial/life cycle, socioeconomic, and heritage were each subdivided by the generation dimension into measures of the parental generation and current characteristics. Measures of familial/life cycle attributes for the family of origin were: (1) family structure (parents present or absent), (2) number of siblings, and (3) knowledge about parents (whether their educational level was known or not). For the present they are respondent's (1) age, (2) sex, (3) marital status, and (4) parental status. On the SES dimension measures for the family of origin were: (1) education of parents, (2) father's occupational prestige, and (3) family income level, and for the present, respondent's: (1) education, (2) occupational prestige, and (3) family income. The heritage dimension measured the cultural and historical background. For the parental generation it asked: (1) region, (2) community type, (3) religion raised in, and (4) farm versus non-farm occupation of father. For the present four parallel measures were used.[8]

Looking at the parental family attributes, Table 4 shows that being high on ethnic identification was associated with having parents present, lots of siblings, and knowledge of mother and father's education. Current family and life cycle attributes, on the other hand, proved to have no association with ethnic identification. On the SES dimension, for both parental and current measures being high on education, income, or occupational prestige associates with having high ethnic identification. On the heritage variable both for the parental and current measures high ethnic identification is associated with being non-southern, urban, non-Protestant (and non-Baptist with Protestant denominations), and non-farm occupations.

In sum, SES and heritage are both associated with ethnic identification and the associations are quite similar for both the parental and current generations. Familial/life cycle attributes, on the other hand, are only

TABLE 4

ASSOCIATES OF ETHNIC IDENTIFICATION GSS: 1972–1977 (WHITES)

	Probability[a]	γ
I. Familial/life cycle		
A. Parental		
Family structure	sig.	.088
Siblings	sig.	.201
Knowledge	sig.	.127
B. Current		
Age	not	—
Sex	not	—
Marital	not	—
Children	not	—
II. Socioeconomic level		
A. Parental		
Mother's education	sig.	.288
Father's education	sig.	.249
Father's occupation prestige	sig.	.115
Family income level	sig.	.120
B. Current		
Education	sig.	.299
Occupational prestige	sig.	.150
Family income	sig.	.159
III. Heritage		
A. Parental		
Region	sig.	.452
Community type	sig.	.499
Religion	sig.	−.231
denomination	sig.	−.363
Industry	sig.	−.311
B. Current		
Region	sig.	.418
Community type	sig.	.259
Religion	sig.	−.221
denomination	sig.	−.432
Industry	sig.	−.091

[a] Statistical significance at .05 level.

associated for the parental generation and none of the current measures are related to ethnic identification.

In order to reduce further this large number of variables to a manageable number suitable for multivariate analysis, all of the variables with parallel measures for the parental generations and the present were run together with ethnic identification. For example, current family income was cross-tabulated with ethnic identification with controls for parental family income. This revealed that (1) in one case the current attribute

(farm/non-farm occupation) was not independently related to ethnic identification, (2) parental and current income, residence, occupational prestige, and education were each independently related, and (3) parental and current religion/denomination and region were so highly interrelated to raise multicolinearity problems. On the basis of these findings, current farm/non-farm occupation was dropped from the analysis and religion/denomination and region scales were made from the parental and current variables.[9] In addition, mother's and father's education were combined into a single measure of parental education (average education of both was used) and knowledge about parents and family structure were combined into a single scale.[10] These procedures reduced the original 24 variables down to 13 variables: number of siblings and parental information; parent's education, father's occupational prestige, and parental income; respondent's education, income, and occupational prestige; parental industry and residence; respondent's residence; and combined parental/current region and parental/current religion/denomination.

At this stage one more problem remained. Parental information and parent's education could not be used together since people eliminated from the analysis because they did not know their parent's education formed an important part of the parental information variable. When preliminary analysis indicated that both variables remained significant explainers of ethnic identification when used separately, it was decided to create a combined variable with those lacking information on one or more parents forming the first category, and those with low (less than high school), medium (high school), and high (some college) education forming the other three categories.

In Table 5 these variables were run with naming a country versus naming no country. The heritage variables clearly work the best. Region, religion, father's industry, and current residence are all related to identification. Only parental residence is not significantly related. SES shows more mixed results. Respondent's education and prestige are related but income is not. Similarly, parent's income is related but father's occupational prestige is not. The parental information/education variable is also related suggesting that both parts of this variable are associated. The other family characteristic, number of siblings, however, is not related.

It thus appears that being unable to name any country is related to a heritage that can be described as Southern, rural, and English–Protestant.[11] This can in turn be described as an old stock, host culture consisting of principally British groups (English, Scottish, Welsh, and Scotch–Irish) that immigrated to America prior to the middle of the 19th century.

Unfortunately, our heritage variables do not permit us to distinguish whether it is immigration generation, affiliation with the host culture, or

TABLE 5

MULTIPLE STEPWISE REGRESSION ANALYSIS OF ETHNIC IDENTIFICATON
(Names Country vs Names No Country)

	Standardized coefficient	F
Variables in Equation		
Region (southern origin vs other)	−.214	228.9
Religion (English–Protestant origins and current vs others)	−.150	113.3
Education (years of schooling)	−.059	17.4
Prestige (prestige score)	−.066	22.4
Father's industry (farm vs non-farm)	−.050	13.4
Current residence (rural vs urban)	−.047	12.0
Parental information/education (don't know, low, medium vs high)	−.098	25.7
High parental income (low, medium vs high)	−.029	4.7
Parental information/education (don't know, low, high vs medium)	−.083	18.6
Parental information/Education (don't know, medium, high vs low)	−0.85	15.7
Variables not in equation		
Income (income level)	−.015	1.2
Medium parental income (medium, low vs high)	−.011	0.5
Father's prestige (prestige score)	−.011	0.6
Parental residence (rural vs urban)	−.000	0.0

both that causes the association between heritage and identification. Evidence from other sources, however, suggests that both factors probably play a role. A review of empirical evidence on assimilation, ethnic identity, and generation suggests that a decline in ethnic identification does occur with generation. This relationship appears in studies by Abramson (1973, pp. 77, 98), Masuda, et al. (1970, p. 203), Sanberg (1974, p. 203), Zeman (1973; pp. 30–44), and Obidinski (1968). Only Pavlak's equivocal findings (1976, p. 29) diverge from the consensus. In general, it appears from evidence at hand that both ethnic homogeneity and identification decreases with generation. This does not mean that ethnicity is dying out as the melting pot theory suggests, since an ample level of homogeneity and identification remains and because there is little evidence as to whether the decline continues beyond the third generation. Still, on the basis of this evidence, it is reasonable to suppose that with generational residence in the United States, ethnic homogeneity and identification does decrease (with declines in each probably leading to further declines of the other) and as a result people tend to either lose track of their ethnicity or gather so many ancestral national lineages as to obliterate any particular ethnic identity.

Additionally, there is evidence that membership in the host culture is associated with low identification. ~~Arnold~~ Driedger found that among college students at the University of Manitoba, those of British background ranked sixth out of seven ethnic groups on both attitudinal and behavioral measures of ethnic identification, ranked seventh on ideal ethnic affirmation, and tied for fifth/sixth on real affirmation. In brief, by all measures, their ethnic identification was comparatively weak. On ethnic denial, however, they ranked last on ideal denial and sixth on real denial. This indicated that they were not trying to suppress their ethnicity, but rather simply lacked much identification of this type (Driedger, 1975, 1976).

Second, nonethnics tend to score low on a number of the SES measures tested here. In particular, low education during the parental and current generation seems to be related to lacking an ethnicity. This can be interpreted to mean that better educated and better-situated family lines are more successful in passing ethnic (and probably other) information along to their descendents than lower-educated and lower-positioned families. This corresponds with evidence from political science that the less educated and those of lower SES have less knowledge about the identity of office holders and the organization of government (Gallup, 1972, pp. 1216–1218, 1378–1379, 1429, 1472, 1606; 1978, pp. 684, 1178) and from survey methodology which finds low education related to giving "Don't know" responses (Sudman & Bradburn, 1974). The ethnic data expand on these finds and indicate that low education and standing are not only related to a low level of knowledge on public facts, but also on personal factual matters.

Finally, the intergeneration transmission of information appears to be an important factor. Families with either a parent or parents absent or with a limited flow of information between generations are less likely to report an ethnicity.

In sum, the prime measurable reasons for not being able to name a country are (1) being part of the old stock, host culture, (2) having low education and standing, and (3) having a break in the intergenerational transmission of information between themselves and parents or possibly between some earlier generations.

Next, in Table 6, the same variables were used to examine the correlates of naming a country versus naming multiple ethnicities. Few of the variables are related to this difference. One measure from each dimension is significantly related, but even these few relationships are quite small. For what it is worth, the analysis suggests that those with too much ethnicity tend to be English–Protestants, come from high prestige families, and have fewer siblings. None of these are considered to tap the fundamental reasons for having too much ethnicity.

TABLE 6

MULTIPLE STEPWISE REGRESSION ANALYSIS OF ETHNIC IDENTIFICATION
(Names Country vs Names Multiple Countries)

	Standardized coefficient	F
Variables in equation		
Religion (English–Protestant origin/current vs others)	–.071	27.0
Father's occupational prestige (prestige score)	.039	8.3
Siblings (number)	.036	6.8
Variables not in equation		
Parental information/education 2	–.018	1.5
Parental information/education 3	–.005	0.1
Parental information/education 4	007	0.2
Education (years of schooling)	.020	2.0
Income (income level)	.002	0.0
Occupational prestige (prestige score)	–.008	0.3
High income	–.020	2.1
Medium income	–.010	0.5
Father's industry	.017	1.5
Region	.020	1.9
Current residence	.018	1.8
Parental residence	–.005	0.1

It was considered that the variables failed to differentiate because the names one country category contained a number of people who originally named more than one country, like the multiple ethnics, but were able to choose one nationality. Perhaps naming multiple ethnicities was the important feature and being able to choose one from among them was not crucial. To look at this feature, those with too much ethnicity were compared only to those naming a single country (and thereby excluding those who named multiple ethnicities and then chose one prime ethnicity). This comparison picked up three more variables significantly related to ethnic identification (father's industry, respondent's education, and being high on parental information/education), but did not reveal any basic difference. These were all only weakly related (β/F = .036/4.8, .037/4.4, .034/4.2). These findings suggest that having too much ethnicity is weakly related to being in America sufficiently long to have picked up more than one nationality in ancestral lines and being of higher educational standing so that the information was passed along. The problem with describing this difference is that the major reason for giving multiple ethnicities are (1) having more than one known national lineage and (2) being unable to choose between them because the mixture is equal (e.g., parents of differ-

ent national origins) or so mixed that no single affiliation is meaningful (e.g., your typical English–German–Irish–Ukrainian) and there are no measures of these characteristics. Nor are there any good stand-ins. Mixed ancestry is probably related to the number of generations in the United States since most immigrants marry within their national group and it is only over several generations that other nationality lines are introduced as the slow melting occurs (Abramson, 1973). On the other hand, number of generations in the United States is also related to less ethnicity (i.e., the country has become more mixed across time) and to living in areas where there was less opportunity for mixed marriages because there were fewer different groups to mix together.

What have we learned from the preceding discussion? First, ethnicity is a difficult attribute to measure, but by using a simple subjective question it can be determined for about 78% of all non-blacks and with a combination of subjective and natal measures can be found for about 88%. Second, ethnic identification does not start with given nationalities, but rather begins with whether a person has an ethnicity to identify with. Third, a lack of any ethnic affiliation is associated with three main factors: (1) being a member of the old stock, host culture, (2) having low education and social standing, and (3) poor transmission of family information between generations. Fourth, ethnic over-affiliation is harder to explain with the data at hand, but is probably most related to mixed lineage, and also probably related to higher education and standing (being more likely to pass on and retain complex lineage information) and immigrant generation (which is related to mixed lineage).

The implication of these findings is twofold. First, work on ethnic identification should be reappraised to take the matter of the "nonethnics" into account. Second, work on ethnic groups should consider the utility of the various means of ethnic measurement available and also take the "nonethnics" into account in their analysis of ethnic differences in American society.

NOTES

[1] For the purposes of this paper ethnicity is used primarily as a reference to nationality or cultural origin. This does not preclude the use of religion, race, and language from the specification of ethnic groups but rather reflects this paper's emphasis on the nationality aspect of ethnicity.

[2] In addition to these there are several other methods of identifying national origins such as the surname approach or the physical characteristics approach. Neither these nor other techniques are generally reliable or commonly employed.

[3] In actuality, most users of list samples do not proport that their list includes all members of a given nationality in a given area or that persons not listed can not be members of the nationality. In practice, however, the list does define the universe to be sampled. It should be further emphasized that not all list samples of nationalities are examples of the behavioral approach. Membership lists fit the behavioral mold since they certify a certain behavior

(membership in a certain organization). Such lists as nationality directories often do not since they are commonly compiled not only from membership rosters, but also from surname searches of telephone directories and personal associations.

[4] In this and the following discussion of the three principal approaches to ethnic identification, the focus is on the major pluses and minuses. In the discussion of the nativity approach, the problems of nonlinear national origins ascent (such as all four grandparents being born in the United States, but one or more parents being born outside the United States); births at sea, in a country during transient, or in a country as a foreign alien (e.g., the children of foreign service officers born overseas); the definition of country; the handling of adopted versus biological ancestors; and other matters are not dealt with.

[5] Both because of mortality and associational patterns (i.e., with each prior generation the ancestor is less likely to be alive during the respondent's lifetime and if alive is likely to have less contact with the respondent), knowledge about ancestors falls off quickly with each intervening generation. A NORC mobility study found, for example, that while 98% of respondents knew their father's occupation, only 76% knew the occupation of their paternal grandfather. Knowledge about paternal great grandfather could be expected to decline as sharply. Also, in their study of kinship, David M. Schneider and Calvert B. Cottrell found that while 55% of their white, middle-class Chicago sample could give the first or last names of all four grandparents, only 14% could identify one-half (four) of their great grandparents (Schneider & Cottrell, 1975, pp. 65–66).

[6] This version was used by the CPS in 1973 and with modification in 1971–1972 and 1974–1977.

[7] This dimension is formally distinct from national origins and as a result has not been discussed previously. In practice, as we will see, immigration generation is related both to ethnic identification and to particular ethnicities as well.

[8] For details on all the measures used, see Davis et al. (1978).

[9] From an extensive analysis of the religion and denomination variables the following groups were isolated: (1) "English" Protestants and those with no religion, (2) non-English Protestants (chiefly German and including Brethren, Moravian, Dutch Reform, Evangelical Reform, Mennonite, Reform, and Lutheran), and non-Protestants (Catholics, Jews, and Others). The non-English Protestants and non-Protestants were then grouped together and their religion of origin and current religion were compared. This gave three groups (1) those English–Protestant at both times, (2) those of changing affiliation, and (3) those not English–Protestant at both times. Eventually to simplify matters further a dichotomy of English–Protestants at both times versus others was used.

[10] If a person did not live with a parent, he was not asked about that person's education (or other characteristics). As a result, for each parent there are three possibilities: (1) that the parent was absent, (2) that the parent was present and their education was not known, and (3) that the parent was present and their education was known. As the figures below show, naming an ethnicity was related to knowing parental education but there were no differences between a parent being absent or not have knowledge about the parent's education and naming an ethnicity.

M = A F = A	M = A F = DK	M = A F = ED	M = DK F = A	M = DK F = DK	M = DK F = ED	M = ED F = A	M = ED F = DK	M = ED F = ED
Naming no country:								
.208 (77)	.274 (62)	.174 (202)	.279 (190)	.220 (869)	.170 (282)	.141 (821)	.166 (492)	.103 (6064)

M = Mother, F = Father, A = Absent, ED = knows education, DK = does not know education.

For example, the 20.8% naming no country among those with both parents absent was quite

similar to the 22% among those with both parents but not knowing their education. Interestingly it appears that a combination of not having one present and not knowing the education of the other produces the least identification. Since it appears that both absence and lack of knowledge were measures of a lack of intergenerational transference of information and both had a similar impact on knowing an ethnicity, they were combined into a single measure of information about parents.

[11] The old stock, host groups are concentrated in the South because most later immigrants settled outside the South, thereby not diluting the original British stock. In addition, white/ black racial divisions may have heightened the sense of "whiteness" and led to the deemphasis of ethnic differences among whites.

REFERENCES

Abramson, H. J. (1973), *Ethnic Diversity in Catholic America*. New York, Wiley.

Abramson, H. J. (1975), The religioethnic factor and the American experience: Another look at the three-generations hypothesis. *Ethnicity* 2, 163–177.

Barton, J. J. (1975), *Peasants and Strangers: Italians, Rumanians, and Slovaks in an American City, 1890–1950*. Cambridge, MA, Harvard Univ. Press.

Cohen, A. (Ed.) (1974), Introduction: The lesson of ethnicity, in *Urban Ethnicity*. London: Tavistock.

Dashefsky, A. (1972), And the search goes on: The meaning of religio-ethnic identity and identification. *Sociological Analysis* 33, 239–245.

Dashefsky, A. (1975), Theoretical frameworks in the study of ethnic identity: Towards a social psychology of ethnicity. *Ethnicity*, 2, 10–18.

Davis, J. A., Smith T. W., and Stephenson, C. B. (1978), *General Social Surveys, 1972–1978: Cumulative Codebooks*. Chicago: NORC.

Driedger, L. (1975), In search of cultural identity factors: A comparison of ethnic students. *Canadian Review of Sociology and Anthropology* 12, 150–162.

Driedger, L. (1976), Ethnic self-identity: A comparison of ingroup Evaluation. *Sociometry* 39, 131–141.

Gallup, G. H. (1978), *The Gallup Poll: Public Opinion, 1972–1977*. Wilmington, DE, Scholarly Resources.

Gallup, G. H. (1972), *The Gallup Poll: Public Opinion, 1935–1971*. New York: Random House.

Glaser, D. (1958), Dynamics of ethnic identification. *American Sociological Review* 23, 31–40.

Gordon, M. (1964), *Assimilation in American Life*. New York, Oxford Univ. Press.

Greeley, A. M. (1974), *Ethnicity in the United States: A Preliminary Reconnaissance*, Wiley Series in Urban Research. New York, Wiley.

Isajiw, W. W. (1974), Definitions of ethnicity. *Ethnicity* 1, 111–124.

Keyes, C. F. (1976), Towards a new formulation of the concept of ethnic group. *Ethnicity* 3, 202–213.

Masuda, M., Matsumoto, G. H., and Meredith, G. M. Ethnic identity in three generations of Japanese Americans. *Journal of Social Psychology* 81, 199–207.

Nahirny, V. and Fishman, J. A. (1976), American immigrant groups: Ethnic identification and the problem of generations. *The Sociological Review* 13, 311–326.

Obidinski, E. (1968), *Ethnic to Status Group: A Study of Polish Americans in Buffalo*. Unpublished Ph.D. dissertation, State University of New York.

Parenti, M. (1967), Ethnic politics and the persistence of ethnic identification. *American Political Science Review* 61, 717–726.

Pavlak, T. J. (1976), *Ethnic Identification and Political Behavior*. San Francisco, R and E Research Associates.

Plax, M. (1972), On studying ethnicity. *Public Opinion Quarterly* 36, 99–104.

Sandberg, N. C. (1974), *Ethnic Identity and Assimilation: The Polish–American Community, Case Study of Metropolitan Los Angeles*. New York, Praeger.

Schneider, D. M., and Cottrell, C. B. (1975), *The American Kin Universe: A Genealogical Study*. Chicago, Univ. of Chicago.

Segelman, R. (1967), Jewish identity scales: A report. *Jewish Social Studies* 71, 92–111.

Sudman, S. and Bradburn, N. A. (1974), *Response Effects in Survey: A Review and Synthesis*. Chicago, Aldine.

Uyeki, E. S. (1976), Correlates of ethnic identification. *American Sociological Review* 41, 391–403.

Vrga, D. J. (1971), Serbians in America: The effects of status discrepancy on ethno-religious factionalism, in *Ethnic Groups in the City* (O. Feinstein, Ed.), Lexington, MA Heath.

Zeman, D. J. (1973), *E Pluribus Unum?: A Study of Four Catholic Ethnic Groups in Chicago*. Unpublished Ph.D. dissertation, University of Chicago.