PROBLEMS IN ETHNIC MEASUREMENT:

OVER-, UNDER- AND MISIDENTIFICATION

GSS TECHNICAL REPORT NO. 29

FEBRUARY, 1982

TOM W. SMITH
SENIOR STUDY DIRECTOR
NATIONAL OPINION RESEARCH CENTER

Over the last decade and a half there has been a great increase in scholarly and governmental interest in ethnicity. Several interdisciplinary journals have appeared (Journal of Ethnic Studies, 1973-; Ethnicity, 1974-; Ethnic Groups, 1978-; Journal of American Ethnic History, 1981-) and the U.S. Census included a general ethnicity question for the first time in 1980. Both scholars and the government have found that ethnicity is a difficult concept to define and measure. There is some general concensus that ethnicity represents a cultural heritage or identification that is specified by some combination of nationality, language, religion, and race (Isajiw, 1974). There is less agreement on how each of these tracers should be measured or combined.

Most approaches favor separate inquires about each of the major components—nationality, language, race, and religion. (Although the Census feels constitutionally prohibited from including the last element.) Other approaches have tried to craft a single item to measure ethnicity, although additional questions on other components may be included.

While measurement problems have arisen for each component, the largest difficulties have arisen from the general nationality measures.[1] Three prominant techniques (the natal, behavioral, and subjective) and combinations of these have been used to measure nationality. All techniques have certain benefits and supply certain unique information. The subjective and quasi-subjective techniques have gained increased acceptance over the last decade and are now the prime measures of ethnicity in most governmental and scholarly surveys. The acceptance of subjective ethnicity measures has resulted from

---

[1]For example, on language see Bureau of the Census, 1974. Racial issues have concerned what groups are considered as "races" and whether the code should be by observation or direct inquiry. Religious questions have concerned themselves with the coding and distinguishing of various sects and denominations and other issues - e.g. McCourt and Taylor, 1976.

their ability to handle more effectively the three major problems of ethnic
measurement; overidentification—mentioning two or more distinct ethnicities,
underidentification—mentioning no ethnicity, and misidentifications—
mentioning an incorrect or inappropriate ethnicity (For comparisons with the
natal and behavioral, see Smith, 1980 and Smith, 1981).

In this paper we will examine these problems and how over-, under-,
and misidentification are handled by subjective ethnicity measures. We will
secondly consider whether the ability of subjective ethnicity measures to
reduce overidentification and minimize underidentification might be more
apparent than real. We will consider whether subjective ethnicity measures
create more identifications, but less reliable measurements. Finally, we will
consider what the data on ethnic measurement tells about assimulation and
pluralism in contemporary and future American society.

As Table 1 reveals both over- and underidentification are major
problems even for subjective measures of ethnicity. A minimum of about 10
percent of non-blacks cannot name any national origin.[2] They either reply
"don't know," "none," "American," or some similar response to ethnic
measures. At the other extreme slightly over 10 percent report two or more
ethnicities and cannot select a main ethnicity. That leaves only 75-80
percent with an ethnic identity. Among these only about two-thirds (or 50
percent of all) have a single ethnicity. The others select a main ethnicity
from competing heritages.

Both between and within these groups the strength of ethnic identity
varies greatly. The 1978 election survey included a four part question
measuring strength of ethnic identity: "In addition to being American, is

---

[2]Blacks are often excluded from ethnic questions (e.g. SRC7762) and
assigned their race as their ethnicity. When asked their ethnicity most name
Africa or their race while a large number name no ethnicity. See Smith, 1978.

TABLE 1

ETHNIC IDENTIFICAL (NON-BLACKS)

	Elec78	GSS72-80	SRC7762
Underidentification:..............	.160	.103	.102
Indentification:			
Names One Ethnicity.............	.463 .738[a]	.520 .780	.846[b] .846
Choses One Main Ethnicity........	.263	.260	--
Overidentification:			
Names two or More Ethnicities:			
Can't Select...................	.102	.117	.052
	(2048)	(10688)	(3278)

[a]Includes .012 from a miscellaneous group.

[b]The high take rate and the elimination of the chose category comes from the use of broad regional categories when particular ethnicities are not codeable. For example, the following codes were used: Europe = .027; Inter-European regions=.103; other regions such as Balkans, Eastern Europe, etc. = .063. Subtracting these regions we find .653 naming one ethnicity.

Wordings:

Elec78: Were both your parents born in this country?
IF NO:
Which country was your father born in?
Which country was your mother born in?
IF YES or DON'T KNOW:
Do you remember which country your family came from originally on your father's side?
Do you remember which country your family came from originally on your mother's side?
IF MORE THAN ONE COUNTRY MENTIONED:
Aside from being American, which country of those you just mentioned do you think of your family as mainly coming from?
Where were you born?

SRC7762: Aside from being American, what nationality do you think of your family being mainly?

GSS: From what countries or part of the world did your ancestors come?
IF MORE THAN ONE COUNTRY NAMED:
Which one of these countries do you feel closer to?

there another nationality or ethnic group you feel you belong to? IF YES:
What group is that? IF MORE THAN ONE COUNTRY NAMED: With which of these
groups do you most closely identify? How strongly do you feel a part of (NAME
OF ETHNIC GROUP)? Would you say you identify very strongly, somewhat strongly
or not strongly?" Less than 2 percent of nonidentifiers indicated that they
belonged to a group and a minuscle 0.6 percent had a very strong attachment.
By contrast 28.4 percent of those with parents from the same background
reported belonging to an ethnic group and 12.9 percent had very strong
attachments. Those with parents from different heritages who were able to
select a main ethnicity reported less identifcation—17.2 percent belonging
and 3.4 percent with very strong attachment. Among those with mixed heritages
who were unable to select a main ethnicity only 8.5 percent reported belonging
to a group. In brief, identification is highest among people with a
homogeneous heritage. It declines for those having to chose between parental
origins or who lack knowledge about one side of their family. Identification
is still lower among those who are unable to select between origins and in
effect the data confirms that those without any reported origins have no
ethnicity to identify with.[3]

Underidentification

Following our previous analysis (Smith, 1980), ethnic identification
was related to four dimensions—familial and life cycle, socioeconomic level,
heritage, and generation. Familial/life cycle, socioeconomic level, and
heritage were each broken down by generation into measures of parental
generation and current condition. Preliminary analysis (Smith, 1980)

---

[3]In addition those who select a group or regional identity (see Table
1, note a) tend to have weaker identifications. While 22.9% of those with
origins from individual ethnicities reported belonging only 12.9% of those
with vaguer, regional origins were belongers.

TABLE 2

STRENGTH OF ETHNIC IDENTITY
(NON-BLACKS)

Information on National Origins	Proportion belonging to Ethnicity	Proportion with very strong Attachment
Names no Ethnicity	.022	.006
Names One Ethnicity		
Paternal or Maternal (one only)	.070	.020
Paternal and Maternal (both the same)	.284	.129
Choses One Main Ethnicity	.172	.034
Names two or more countries;		
Can't Select	.249[a]	.085

Source: Elec78 (1965)

[a]Most of the cases involving people who did not choose a country, but who said they belonged to a group included Amerindians or Jews who were not coded in a comparable fashion on both variables. Excluding these cases the proportion belong falls to .085. For similar reasons the proportion belonging among those naming no ethnicity decreases to .016.

indicated that each of the family/life cycle measures for parents (being raised by parents, remembering information about parents, and number of siblings) were related to ethnic identification, but none of the current variables (age, sex, marital status, and number of children) were related. All SES variables (education, occupational prestige, and income) show associations both for parental and current generation. Parental education was however the source of our parental information measure so we had to combine these measures into a variable with lack of information about parental education the first category, those will less than a high school education, high school graduates, and those with some college forming the next three categories. Heritage variables measure (indirectly) membership in our old, host British culture (region, community type, religion/denomination, and farm/nonfarm employment). They also all showed initial associations with ethnic identification. Current

farm/nonfarm employment did not have an independent effect once father's employment was controlled, however, and the associations between parental and current region and religion were so high that to avoid problems of multicolinearity they were combined into single measures. In addition we were able to add to the previously employed variables a new and somewhat more direct measure of ethnic heritage, immigrant generation (1st generation = foreign born; 2nd generation = native born of foreign born parents; 3rd generation = native born parents and foreign born grandparents; and 4th plus generations = native born of native born parents and grandparents). This resulted in 13 variables: number of siblings; combined parental education/knowledge; parental occupational prestige and income; current occupational prestige, education and income; parental industry and residence; current residence; combined parental/current region and parental/current religion-denomincation; and immigrant generation.

The analysis in Table 3 basically confirms our previous analysis of the causes of underindentification. The heritage variables clearly work the best. Religion, immigrant generation, and region all consistently relate and farm origin and rural residence have weak associations. Only current industry and parental residence are never significantly related. SES has a somewhat more mixed and modest effects. Education, income, and parental income are related, but father's and respondent's occupational prestige are not. Finally, the parental information/education variable is related, while the other family variable, number of siblings, is not.

It thus appears that being unable to name a country is most strongly related to being members of an old stock, host culture consisting mostly of British ethnic groups (English, Scottish, Welsh, and Scotch-Irish) that immigrated to American prior to the middle of the 19th century. These people have ethnic amnesia having forgotten their origins because of both their long residence and association with the majoritarian culture. Secondly, nonethnics

TABLE 3

MULTIPLE STEPWISE REGRESSION ANALYSIS OF ETHNIC IDENTIFICATION
(WHITES)
(NAMES COUNTRY VS. NAMES NO COUNTRY)

Variables in Equation	1972/77 Revised[a]	1977-80	1977-80, Exclusions
	Standardized Coefficient/F	SC/F_1	SC/F
Immigrant Generation	--	.130/38.2	.135/51.8
Region (southern origin vs. other)	-.207/209.5	-.210/102.2	-.205/121.6
Religion (English-Protestant origins and current vs. others)	-.155/119.4	-.112/26.8	-.120/38.6
Education (Years of schooling)	-.154/125.3	-.131/37.3	-.143/55.7
Father's industry (farm vs. non-farm)	-.043/9.5	(-.026/1.9)[b]	--
Current residence (rural vs. urban)	-.042/9.5	(.006/0.1)	(.014/0.7)
High parental income (Low, medium vs. high	-.032/5.8	-.065/7.7	-.033/3.9
Variables Not in Equation			
Occupational Prestige (NORC prestige)	-.029/3.5	-.026/1.4	--
Parental information/education (dont know, low, medium vs. high)	-.027/3.6	(-.052/6.6)	(-.058/10.6)
Father's Prestige (NORC prestige)	-.018/1.6	-.004/0.0	--
Industry (farm vs. non-farm	.013/0.9	.001/0.0	--
Medium parental income (Medium, low vs. high)	-.010/0.5	(-.044/4.0)	-.036/3.7
Parental information/education (dont know, low, high vs. medium)	-.007/0.2	-.008/0.2	-.019/1.2
Siblings (number)	.007/0.3	.039/3.7	.031/2.9
Income (income level)	-.005/0.1	(-.074/13.9)	(-.079/19.7)
Parental information/education (dont know, medium, high vs. low)	.001/0.0	-.009/0.2	-.022/1.4
Parental residence (rural vs. urban)	-.000/0.0	.020/1.0	-.020/1.3
n	5081	2474	3034
r^2	.145	.179	.190

[a]These results differ from those reports in Smith, 1980, p. 89 for
three reasons: 1) certain variables used in the equations were omitted from
the table, 2) the current residence variable used was replaced by a revised
code, and 3) errors in the coding of the education variable were corrected.

[b]Parentheses indicate that a variable listed in the equation group has
dropped out of the equation or that a variable listed as not in the equation
has entered it.

tend to score low on SES measures. This agrees with much other data about the lower knowledge and greater item nonresponse among those with lower SES and indicates that the difference extends to knowledge about one's own family. Finally, nonethnics tend to come from families with either low information about parents or a break in contact with parents.[4]

Overidentification

In our previous analysis of overidentification we indicated that none of the explanatory variables clearly related to the basic causes of overidentification: 1) having more than one known ethnicity and 2) being unable to choose between them. Trying numerous combinations of variables on the 1977-80 and 1980 GSS and the 1978 election survey we found only one notable and consistent explanatory variable: number and pattern of ethnic identities.[5]

---

[4]These factors are confirmed by a similar analysis of the 1978 election study. Not mentioning a country was significantly related to membership in a "Southern" religion (standardized coefficient/f=.170/48.7); being raised in the South (.178/53.9); rural parental residence (.092/14.5); less education (-.208/70.9); not knowing parental information (-.054/5.7), and youth (-.082/11.5). Income and father's work on a farm were not related.

[5]A supplement to the 1980 GSS asked the following additional questions about national origins:

In what country was your (MOTHER/FATHER) born?
IF BORN IN USA:
What countries or parts of the world did your
(MOTHER'S/FATHER'S) ancestors come from? LIST FIRST
TWO MENTIONS.

We used these variables to create two measures of the pattern of ethnic information. First, we counted the total number of distinct ethnicities mentioned, from none to four. Second, among those with two or more ethnicities we coded the differences as a) between parents, b) within parents, c) both between and within parents, and d) within parents with one ethnicity shared between parents (e.g. either Mother: Irish-German and Father : Irish-German or Mother: Irish-French and Father: Irish-French). In essence this measured a potential "overload" of ethnic information. We assumed that the more information a person has and the greater the conflict between contesting ethnicities, the greater the likelihood that no ethnicity would be chosen.

Ethnic Reports for Parents	% Can't choose on Subjective Ethnicity
Number Mentioned:	
None	.045
One	.031
Two+: Point of Clevage:	
Shared ethnicity between parents; difference within parent(s)	.109
Difference between parents	.174
Difference within & between parents	.194
Differences within parents	.300
n = 1429	

Overidentification is practically nonexistent among those who report
no origins for their parents or a single ethnicity. It increases for those
with a shared ethnicity between parents and in higher still when no shared
ethnicity appears. The only surprise is the fact that the greatest degree of
overidentification appears among those with differences within, but not
between parents rather than among those with differences both within and
between parents. Since the number in the former group is small we cannot put
too much weight on this.

We then examined among people mentioning two or more countries why
some were able to select an ethnicity while others were not. First, there
appears to be great instability in being able to select an ethnicity in this
situation and random and transitiory factors undoubtedly play a large role
(see below). Looking at what could be explained beyond the measurement error,
we found that this decision apparently had little to do with ethnic heritage.
Being able to select a main ethnicity was associated with being better
educated, not replying DK to attitude questions, and having children. Having
a shared ethnicity between parents and urban residence were also related to
selecting a main ethnicity, but the associations were of borderline
significance. Religion and age were unrelated and preliminary analysis also
indicated no association with sex, presence of parents at age 16, number of

siblings, region, knowledge about parents, and sureness about maternal and
paternal information. The associations between selecting a country and better
education and firmer attitudes probably indicates that these people are less
likely to be overloaded by the crosscutting identities. They are more
cognitively poised and able to chose one ethnicity among the mixture. The
higher rate of selection among those with parents sharing an identity
indicates that the conflict between differing ethnicities is simplified by
having common origin. The lower selection among rural residents may be a
reflection of their longer presence in America. While aware of certain
national origins, they may not feel close to either of them and therefore are
unable to chose between them. We have no explanation for the association
between higher fertility and being able to select a country.

Misidentification

In previous discussions misidentification was considered a small,
technical problem compared to the large difficulties of over and
underidentification. We had in mind mainly cases where ethnic origin did not
jive with nativity (e.g. the George Romney situation), where multinational
states confused origins (e.g. a Polish ancestor being remembered as Austrian
and Hungarian), or where intentional passing was attempted (as in Kazan's The
Arrangement). (See also Lyman and Douglass, 1973; p. 355.) But when we
discovered that only about 20% of the population said they "belonged" to an
ethnic group, we wondered whether most ethnic identification might be
misidentifications in the sense of being meaningless tags without behavioral
or affective consequences, whether the identification of most people might be
either 1) labile and unreliable (in the classic sense of not yielding
consistent measurement) or 2) reliable but of little meaningful consequence
(in the sense that person may reliably identify himself as having grey-blue
eyes without this classification having any importance to himself or

others). To examine this broad type of misidentification, we studied how the
distribution of ethnicity and the associational pattern of ethnicity varied by
strength of identification. Strength of identification was measured directly
by the belonging-identify scale, by a measure of single and multiple
identifications, and by immigrant generation.

As Table 4 shows the restriction of ethnicity to "belongers"
dramatically changes the distribution of ethnic groups. Recent immigrant
groups make up a much larger share of the "belongers" than among those with
some ethnic identification. Hispanics, Poles, and Italians have their share
doubled among "belongers" as compared to their share on the GSS (the ratios
are similar if the comparison is made between belonging and any other of the
general identification questions). In contrast the old stock, host culture -
English/Welsh - are heavily underrepresented among the "belongers". The
middle stock groups, French, German/Austrian, and Irish, are also
undercounted, but not much as the English/Welsh. The one exception to the
immigrant wave generalization is the Amerindian. While strictly speaking part
of the earliest immigration wave, they are overrepresented among the
"belongers". Unlike old and middle stock Europeans they have not blended into
an American ethnicity thereby losing identification with or even knowledge of
their pre-American origins. Rather they have maintained their cultural
heritage and identification. Yet in another sense they are not really an
exception to the immigrant wave rule. With much of the Amerindian population
still living on reservations and in ethnic communities or just first or second
generation "immigrants" to white society from these ethnic enclaves, it may be
more accurate to think of them as recent immigrants. Other exceptions would
include such groups as the Amish and Hutterites. By living in isolated rural
communities and strictly minimizing intercourse with the main culture these
and similar groups have maintained a separate identity after generations of
residence in the United States. Thus, it appears that it might be degree and

TABLE 4

A Comparison of Ethnic Distributions from Various Surveys

Ethnicity	1978 Election Belong to	1978 Election Combined Items	1978 QOL	1972-80 GSS	1978 Election Belong/GSS
Amerindian	6.1%	1.1%a	2.4b	3.2%	1.91
Hispanic	12.3%	4.4	4.1	4.3	2.86
English/Welsh	3.5	15.8	11.2/15.8c	15.6	0.22
Irish	10.5	12.1	11.6	12.7	0.83
German/Austrian	10.9	23.7	21.1	23.7	0.50
Polish	7.9	4.4	2.6	3.7	2.14
Italian	15.1	6.4	4.9	6.3	2.40
French	1.1	2.7	2.1	2.5	0.44
%with ethnicity	19.9%	68.0%	72.7%	75.6%	

[a] This figure undercounts Amerindians because they were not distinguishable in some codes.

[b] In general the proportions are lower on this survey because all blacks were assigned an ethnicity of "black". On the combined 1978 election items and the GSS many blacks who did not mention an ethnicity were excluded. The higher proportion of blacks in the QOL survey reduces the proportion in 11 other categories.

[c] The low figure includes only "English, British" and "Welsh". The higher figure added in "From British Isles; From two or more countries of British Isles."

or length of _exposure_ to the main American culture that accounts for the falling ratio between proportion among "belongers" and proportion among all identifiers. This correlates highly with immigrant generation for the individual and with timing of immigrant peak among groups. It is also probably conditional on whether racial or religious differences are combined with national origins.

We see from Table 5 that the predictive power of ethnicity declines with the strength of the ethnic identification either measured directly with

TABLE 5

ETHNIC PREDICTIVITY BY STRENGTH OF IDENTITY
(Eta)

Ethnicity[a] by	Ethnic Strength			Generation		
	Belongs, Very Strong	Belong, Somewhat & Not Strong	Doesn't Belong	First	Second	Third+
Region	.321	.341	.212	.358	.313	.215
Religion	.226	.233	.398	.426	.182	.150
Abortion	.292	.210	.149	.276	.274	.130
Women's Rights	.325	.172	.053	.372	.164	.119
ERA	.471	.221	.132	.527	.245	.090
Taxes Wasted	.248	.134	.100	.270	.211	.079
Trust Government	.210	.133	.080	.287	.196	.065
School Integration	.427	.145	.102	.460	.232	.131
Average	.315	.198	.128	.372	.227	.122

[a]Eight ethnic groups were used: Hispanic, English, Irish, Germany, Polish, Italian, Eastern Europe, and others. Source: 1978 election survey.

subjective identification or indirectly with immigrant generation.[6] A similar pattern occurs on the General Social Surveys where the average eta drops from .160 for those naming only one identity to .125 for those naming two or more ethnicities and then choosing a primary affiliation. The decay of ethnicity as strength of identity declines is illustrated by the following figures on religious affiliation for the GSS:

Association between Ethnicity and Religion (ETA)

Names one   Pick One from Two or More

Religion Raised   .436 ⊖ → .323

Present Religion   .350 ⊖ → .244
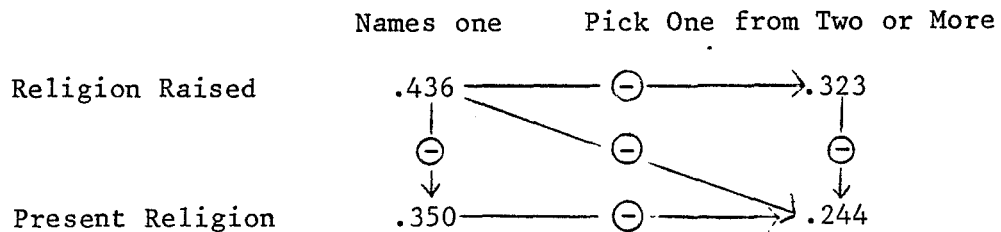
[6]Similar results appeared both when we used a more refined 14 category ethnicity recode and when analysis was restricted to non-blacks.

Whenever we move from strong to weak identification (actually single to mixed)
or from parental generation to present the explanatory power of ethnicity
declines. Similarly when we simultaneously controlled for generation and
belonging on the 1978 election survey we found that both variables have
independent contributions to the explanatory power of ethnicity. Those from
the first or second generations who belonged to a group had an average eta for
the eight variables of .318 while those from the third plus generation who did
not belong to a group had an average eta of .134. Those from the third plus
generation belonging to a group and those from the first or second generations
and not belonging to a group were in betweeen with average etas of
respectively .226 and .246.

Looking at the changes among individual ethnic groups and variables we
see that the declining explanatory power follows certain patterns. For some
variables there is a major mean shift in which all, or almost all, groups
shift in the same direction. For example, the proportion Catholic is lower
for all nationality groups among the fourth plus generation than among the
first/second generation and among those with mixed vs. unmixed ethnicities. A
similar pattern occurs for ideal number of children (proportion favoring three
or more declining). On other variables groups tend to converge, sometimes
with little overall change in the grand mean. Groups high on a variable
tending to lower their levels as identification and generation declines and
groups with initial low levels rising. On government trust for example three
of the four groups above the grand mean for belongers decline among non-
belongers while three of the four low group increased their trust. Similarly
while all four of the most pro-ERA groups have lower support among non-
belongers, two of the four least pro-ERA groups increased their support among
non-belongers. In sum, ethnicity is a more powerful explainer when
identification, whether measured by generation, mixed vs. single identification,

or feelings of belonging, is stronger. As identification weakens, there are sometimes major shifts in distributions for all ethnic groups and sometimes off-setting movement by different groups. In almost every instance however the net result is that explanatory power declines as identification weakens.

Part of the decline comes from the lower reliability of ethnicity among those with weak identifications. Test/retest comparisons on small national subsamples of the 1972, 1973, and 1974 General Social Surveys indicate a notable amount of measurement error in the pattern and degree of ethnic identification. Those reporting one ethnicity at time one were most consistent, 81% again reporting one ethnicity at time two. Consistency in reporting the amount and degree of ethnic identification dropped to 67% for those not naming an ethnicity, 65% for those mentioning two or more ethnicities and selecting between them, and only 48% for those naming two or more ethnicities and unable to chose between them. While committment to a single ethnicity is relatively reliable, underidentification and, to an even greater extent, overidentification are much less reliable. As a result of this unreliability, it will be impossible to explain much of the apparent variance between identifiers, underidentifiers, and overidentifiers since the measurement error will substantially attentuate the true associations.

On the measure of ethnic identification on the GSS we were able to consider whether the lower explanatory power of ethnicity for people with mixed ethnicity (mentioning two, but selecting one) vs. those with a single ethnicity might be the result of more measurement error among the foremore group. The GSS test/retest data indicated that 76.3% of those with a single ethnicity mentioned the same ethnicity at both times while only 69.1% of those with mixed ethnicity. Most of this difference came from greater switching between ethnicities among mixers while there is also a slightly greater chance of mixers not coming up with a main ethnicity at time two. While the greater

unreliability among mixers explains part of the drop in explanatory power, we suspect that even among mixers who consistently select a main ethnicity the explanatory power of ethnicity is probably lower. Even when consistent in their selection of an ethnicity they probably have the explanatory impact of their ethnicity diluted by their other ethnic backgrounds. Similarly, among reliable single identifiers we suspect that cultural impact is less among weak identifiers and attitudinal and behavioral associations are also attenuated.

While the reliability and explanatory power of ethnicity declines with strength of identification, it is not fair to count people with nominal to weak ethnic ties as misindentifications. The test/retest crosstabulations are sufficiently high to indicate that random, wild guessing is not excessive (See also Appendix and Smith, 1981, and even among many weak identifiers and fourth generation Americans ethnicity influenced attitudes and behaviors. On the other hand the various differences by strength of identification are large enough to merit close attention and the importance of having some measures of ethnic strength is clear.

ETHNIC PLURALISM AND ASSIMULATION

As strength of ethnic identification declines, ethnic identification becomes less reliable and its explanatory power is attentuated. Declining strength of identification is strongly tied to immigrant generation (although subjective attachment does exercise a separate, independent effect). On the 1978 election study we find that all respondents with no ethnic identification are from the third+ generation while only 51% of those with very strong attachment are from the third+ generation. Similarly, on the GSSs 97% of those whithout any ethnicity were from the fourth+ generation and only 43% of those with one ethnicity were from the fourth+ generation. This indicates that assimulation, either through forgetting one's roots or intermarriage, does occur (see Isajiw, 1981; Isajiw and Makabe, 1981; and Montero, 1981).

It may seem that eventually all shades of ethnicity will blend into an all-
American khaki (Current, 1981). Cultures do amalgamate. We no longer

distinguish between the Angles, Saxons, and Jutes and the Burgundians have

just about disappeared among the French. Yet several factors are likely to

keep ethnicity an important force in American society for the predictable

future. First, many ethnic differences decline slowly enough and are large

enough to insure notable differences for generations to come even if we assume

inevitable attentuation of ethnic identification. Second, new immigrant

waves, although small in comparison to those of many earlier periods continue

to fill America with more "hard core" ethnics. Third, cultural pluralism and

bilingualism have gained strength in the last ten years (Goering, 1971 and

Novak, 191) and this may well help to perpertuate ethnic identity as it has in

Canada[7] and among isolated American groups (Amish, Cajuns, Amerindians).

Fourth, new ethnicities have been formed within America. While old stock

Yankees and Southerners share fairly similar national origins, they have

established major differences in speech, religion, and other attributes. Of

course, there is some argument that these differences too are declining, but

---

[7]The Canadian Census asks, "To what ethnic or cultural group did you
or your ancestor (on the male side) belong on coming to this continent?
Overidentification is eliminated by arbitrarily taking the paternal lineage.
Underidentification appears to be a trivial problem, in 1971 only 0.96 percent
were in the "other and unknown" category. This included both people giving
miscellaneous ethnicities not listed separately, those listing "Canadian" as
their ethnicity, and apparently those without any ethnic identity. We do not
know how many are those with ethnicity not listed elsewhere, but since 12.0
percent reported a language other than English we suspect that a fair share
are among this ethnically identified group. (Many also report religions with
clear ethnic associations such as Mennonite, Ukrainian Catholic, and Greek
Orthodox.) It thus appears that well less than 1 percent of the Canadian
population lacks an ethnic identity, a rate about a tenth the American
counterpart. We speculate that the biculturalism of Canada has perpetuated
ethnic identification among long term members of the British and French host
cultures and has probably also led new immigrants to retain ethnic
identification (or possibly to adopt the identification of the host culture
they assimulate into--usually the British). The central social and political
importance of culture group in Canada has thus largely prevented the ethnic
amnesia that many long term American residents suffer from. (Kralt, 1977 and
Statistics Canada, 197).

substantial evidence of continuing differences indicates no quick demise of these cultural differences. For these and other reasons, we suspect that analysts will find ethnicity an important background variable for the foreseeable future.

CONCLUSIONS

Of all basic background variables, ethnicity is probably the most difficult to measure (Rothwell and Rustemeyer, 1979 and Scherr, 1980). The analyst must piece together different parts of language, religion, race, nationality, and culture to form ethnic groups. He then finds that about a quarter of the pieces are missing due to over- and underidentification. Both over- and underidentification are partially products of American assimulation. The underidentifiers are overwhelmingly from the old, host immigrant stock; so intergrated into America that their origins in the British Isles and other places have been lost. The overidentifiers have mixed ancestry that is either so hetrogeneous or so balanced that no single ethnicity prevails. While we have no direct evidence, we suspect that ethnic mixers probably more quickly succumb to underidentification than the ethnically homogenous (e.g. lose any identification in fewer generations). We have also found that the explanatory power of ethnicity varies notably by strength of identification. While this strength drops with immigrant generation, it continues among many people in the third, fourth, or higher generation. Given the complexities of ethnicity and the intrinsic origin of the problems of over-, under, and misidentification, there is no simple correction for the large amount of measurement error that exists. Yet by continued studies of these problems we can not only come up with better ethnic measures (and different measures better suited for particular research questions), but also learn more about the processes of assimulation and identification.

APPENDIX: INCONSISTENCIES ON THE ETHNIC MEASURES
OF THE 1978 ELECTION STUDY

On Elec78, 1,195 cases gave an ethnic identity but did not belong to
any ethnicity, 652 cases neither gave an ethnicity nor belonged to any group,
371 both gave an ethnic identity and belonged to a group (and of these 40 or
10.8% mentioned different groups), and 86 cases did not mention a national
origin but belonged to a group. In three cases the "belonging" ethnicity
appeared among the mentioned identifications, but was not clearly mentioned as
the chief origin and in one case an imprecise code hid a probable agreement.
Subtracting these four cases where there was no clear contradictions we are
left with 36 disagreements. 36.1% involved a code of "Jewish" on "belonging"
which did not appear among the nationality oriented identification question,
25.0% involved switches between culturally similar ethnicities (German to
Austrian; British Isles to English, etc.), 13.9% involved large and
inexplicable differences (Eastern Europe-Irish/Amerindian-Irish, etc.), the
final 25.0% inlcude cases of possible national minorities (Germans from
Russia, Poles from Austro-Hungary) and other possible confoundings. The 86
cases where people belonged to a group but did not mention any national
origins were mostly mechanical and artificial differences 43 cases involved
Amerindians who were not distinguishable in the codes used on certain
variables, 20 were people who only mentioned a race as the group they belonged
to, 3 were Jews, and in one case an undefined code appeared. Excluding the 67
coding problems and the racial and religious codes that did not imply any
knowledge about national origins we find that there were 19 discrepancies. Of
these a person contradicted their belonging choice by not selecting this group
as their chief national origin (although mentioning it) in 4 cases, in 6 cases
only other origins were mentioned, and in 9 cases no origins were mentioned.

In many of these cases it appears the failure to come up with the belonging identity was the result of an interviewer's failure either to administer part of the national origin questions or to recode the respondent's response. In brief, we find that the disagreements do not represent the random unreliability of either ethnic measure as they do testify to the complexity of many individual ethnic heritiges.

BIBLIOGRAPHY

Current, Richard N., "The 'New Ethnicity' and American History: Wisconsin as
        a Test Case," The History Teacher, 15 (Nov. 1981), 43-

Goering, John M., "The Emergence of Ethnic Interests-A Case of Serendipity,"
        Social Forces, 49(,1971), 379-384.

Isajiw, Wsevold W., "Definitions of Ethnicity," Ethnicity, I (July, 1974),
        111-124.

Isajiw, Wsevolod W., "Ethnic Identity Retention," Ethnic Pluralism Paper No.
        5, Center for Urban and Community Studies, University of Toronto,
        July, 1981.

Isajiw, Wsevolod W. and Makabe, Tomoko, "Ethnic Identity Retention and
        Socialiation," Paper presented to the American Sociological
        Association, Toronto, August, 1981.

Kralt, John, Ethnic Origins of Canadians. 1971 Census of Canada. Statistics
        Canada. Catalogue 99-709. Vol.-V-Part 1. Bulletin 5.1-9, May, 1977.

Lyman, Stanford M. and Douglass, William A., "Ethnicity: Strategies of
        Collective and Individual Impression Management," Social Research, 40
        (Summer, 1973), 344-365.

McCourt, Kathleen and Taylor D. Garth, "Determining Religious Affiliation
        through Survey Research: A Methodological Note," Public Opinion
        Quarterly, 40 (Spring, 1976), 124-127.

Montero, Darrel, "The Japanese Americans: Changing Patterns of Assimulation
        Over Three Generations," American Sociological Review, 46 (Dec.,
        1981), 829-839.

Novak, Michael, The Rise of the Unmeltable Ethnics: Politics and Culture in
        the Seventies. New York: MacMillan, 1971.

Rothwell, Naomi D. and Rustemeyer, Anitra M., "Studies of Census Mail
        Questionnaires," Journal of Marketing Research, 16 (Aug., 1979),
        401-409.

Scherr, Marrin G., comp. by, "Research to Improve the Quality of Race/Ethnic
        Data Collected by the Federal Government," A Collection of Papers
        Presented at the American Statistical Association, Houston, Aug.,
        1980.

Smith, Tom W., "Ethnic Measurement and Identification", Ethnicity, 7 (March,
        1980), 78-95.

Smith, Tom W., "Ethnic Measurement and Identity," GSS Technical Report No.
        8. Chicago: National Opinion Research Center, 1978.

Smith, Tom W., "The Subjectivity of Ethnicity," Paper prepared for the Panel on Survey Measures of Subjective Phenomena, National Academy of Sciences, April, 1981.

Statistics Canada. 1971 Census of Canada, Population, Ethnic Groups. Bul. 1-3.2, Catalogue 92-723, 197   .

U.S. Bureau of the Census, Accuracy of Data for Selected Population Characteristics as Measured by Reinterviews. PHC (E)-9. Washington, D.C.: GPO, 1974.