MULTIPLE CLASSIFICATION ANALYSIS (MCA): PURPOSE AND AN EXAMPLE

John P. Robinson
University of Maryland

Most surveys involve a very large number of both dependent and predictor variables. As an example, in the first national Survey of Public Participation in the Arts (SPPA) conducted by the U.S. Census Bureau in 1982, there were life-style participation items, media participation items, socialization items, participation barrier items, and music preference items. In addition there were more than 20 predictor variables, manly concerning each respondent=s general social and demographic background.

Cross-tabulation is the most straightforward and traditional way of showing the interrelation of such items in a social survey like the 1984 SPPA. However, in the case of the many variables in most surveys, that would involve several thousand cross-tabulations B an extraordinary number both to process and to display. Moreover, it is an inordinate number to comprehend or to put into larger perspective, particularly since many of the arts participation variables are closely or subtly tied together, (e.g., attending concerts is related to attending the ballet; education is tied to annual income and geographical area).

To put these data in a clearer and broader perspective, a statistical technique called Multiple Classification Analysis (MCA) was developed by Andrews et al. (1973). MCA was developed to provide efficient analysis of multiple variable data sets like that collected in the SPPA. It first shows the interrelations between variables as in a single cross-tabulation, but then further shows the results of cross-tabulations with many other related variables at the same time. It can convey the same incisive conclusions as multiple regression analysis or analysis-of-variance (ANOVA) techniques, but in a way that can be easily comprehended by anyone familiar with the basic logic of a cross-tabulation.

An Example:

The example in Tables 1 to 3 illustrates the logic and value of MCA. The example uses attendance at musicals from the SPPA as the variable to be predicted (called the Adependent@ variable), and it examines how well the respondent=s education and race (the Aindependent@ or predictor variables) can be used to predict such attendance. First it can be seen that some 19% of the respondents in the 1985 SPPA reported attending a musical stage play in the previous 12 months.

The cross-tabulation of attending musicals by education as given in Table 1 shows that such attendance varies widely by that factor: only 4% of those with a grade school education reported attending a musical (i.e., 96% did not attend) and only 6% with some high school education attended. Yet attendance was about 7-10 times as high (45%) at the other end of the education spectrum B those with graduate school education. In other words, respondents with

graduate school education are almost 7 times as likely to say they attend musicals (45%) than those with less than a high school degree (about 5%).

Table 1: Rate of Attendance at Musicals by Education and Race: 1982 SPPA

a) Education:			
	Attend	Not Attend	Total
Grade School (2,067)	4% +	96% =	100%
Some High School (2,238)	6	94	100
High School (6,494)	14	86	100
Some College (3,348)	27	73	100
College Graduate (1,795)	37	63	100
Graduate School (1,300)	45	55	100
Total Sample*	19% +	81% =	100%

b) Race:			
	Attend	Not Attend	Total
White (15,167)	20%	80%	100%
Black (1,673)	10	90	100
Other (403)	13	87	100
Total Sample*	19% +	81% =	100%

*Total sample responses do not sum to exactly 17,254 due to a small number of respondents who failed to answer the education or race items.

There are also large differences in attending musicals by race, as shown in the bottom (second) cross-tabulation in Table 1. Some 20% of all white respondents reported attending musicals in the last 12 months, compared to 10% among blacks and 13% among respondents of other minority racial backgrounds (Asian Americans, Hispanics, etc.). Thus, white attendance exceeded black attendance by 10 percentage points and Aother@ racial group attendance by 7 points.

Table 1, then, contains two independent cross-tabulations, one for education and one for race. However, the two predictor variables of race and education are not independent of each another. Both blacks and many other minority racial groups in the country have less formal education than the white population. That raises the question of how much of the racial differences in attending musicals are basically tied to race vs. how much these are a by-product of the educational differences that exist across these racial groups.

One way to examine this question is to cross-tabulate attendance at musicals by education separately for each racial group. These separate tabulations are shown in the middle columns of Table 2. The first column in parentheses shows the same overall differences by education presented in Table 1. The second column shows these same differences B but only for the white respondents in the survey. The third column shows results only for black respondents and the fourth column only for respondents of Aother@ racial backgrounds. Note the percentages at the bottom of each of the last three columns: they equal the 20% white attendance, 10% black attendance and 13% Aother@ attendance, found in Table 1.

The racial comparisons between columns 2 and 3 of Table 2 (shown in Column 5) are now more precise because they contrast whites and blacks with the same level of education. Grade school educated blacks are compared directly with grade school educated whites, grade school educated persons in Aother@ racial groups, and so on for each educational level.

As might be expected, the overall racial differences of ten percentage points between whites and blacks is reduced considerably within most categories of education. In column 5 grade school educated blacks attended musicals compared to 4% of grade school educated whites, a difference of only 1 point, not 10 points. Similarly for high school graduates, the difference is only 5 percentage points, and not 10 points. The racial differences for the six education groups in order in the next Adifference@ column are 1, 4, 5, 9, 10 and 13 percentage points. These differences average 6 percentage points after weighting for the different size of each educational group (shown in column 6 of Table 2) and added together at the bottom of Table 7.

Table 2: Calculation of Education-Adjusted Differences in White and Black Rates of Attendance at Musicals

Education:	Total Sample (1)	Attended Musicals (2) White	(3) Black	(4) Other	Differences (5) White-Black		6) % of Sample		7) Education Adjusted Differences
Grade School	(4%)	4%	3%	4%	+1	X	12%	=	.12
Some High School	(6%)	7	3	0	+4	X	13%	=	.52
High School Grad	(14%)	14	9	8	+5	X	38%	=	1.90
Some College	(27%)	28	19	16	+9	X	19%	=	1.71
College Graduate	(37%)	39	29	20	+10	X	10%	=	1.00
Grad School	(45%)	46	33	22	+13	X	08%	=	1.04
Total	(19%)	20%	10%	13%			100%	=	6.29

In other words, when we take the step of comparing racial groups with the same educational level, the original 10 point gap between whites and blacks in Table 1 reduces to an average of only 6 points. (When calculated the same way, the difference between whites and Aother@ races increases rather than decreases, with whites averaging a 10% higher participation than persons in other racial groups when education is controlled for, compared to the original, unadjusted difference of 7%).

That is the same analytical logic and approach that is employed in Multiple Classification Analysis. While the MCA procedure does not show the inner details of the separate Table 2 breakdowns, it does show the same end results as shown in Table 3 B the 10 percentage point difference (20% white vs. 10% black) in Table 1 is reduced to an average of 6 points (20% white vs. 14% black) after controlling for differing educational levels among blacks and whites. SPSS results of MCA analyses are generally presented in the format appearing in Table 3. (The SPSS version of MCA can be found in its ANOVA package and does require using the syntax options)

The analysis in Table 2 represents a very simple application of MCA to only two variables (race and education); using the same approach as in Table 2, the reader can calculate the 9 point lower figure for the other race category that emerges in Table 3. However, the world of arts participation and attendance, like other forms of human behavior, involves far more than two or three variables. The unique value of MCA is realized when one uses the technique to separate the effects of not just two but three, five or ten factors that affect participation. For

example, the addition of income and region to the MCA analysis further reduces the differences between whites and blacks.

Table 3: MCA Presentation of Rates of Attendance at Musicals by Education and Race, Before and After MCA Adjustment

	Before Adjustment	After Adjustment
Overall Attendance	19% (Difference from)	19% (Difference from)
Education:		
Grade School	4% (-15)	5% (-14)
Some High School	6 (-13)	6 (-13)
High School	14 (-5)	14 (-5)
Some College	27 (+8)	27 (+8)
College Graduate	37 (+18)	38 (+19)
Graduate School	45 (+26)	45 (+26)
Race:		
White	20% (+1)	20% (+1)
Black	10 (-9)	14 (-5)
Other	13 (-6)	10 (-9)

H:\ROBINSON\2001\MCA.wpd