DOCUMENT RESUME

ED 454 244 TM 032 842

AUTHOR	Witta, E. Lea
TITLE	Does Method of Creating the Sample Influence Missing Data Decisions?
PUB DATE	2001-04-00
NOTE	23p.; Paper presented at the Annual Meeting of the American Educational Research Association (Seattle, WA, April 10-14, 2001).
PUB TYPE	Reports - Evaluative (142) -- Speeches/Meeting Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Decision Making; *High School Students; High Schools; Research Methodology; *Sampling; Student Employment
IDENTIFIERS	*Missing Data

ABSTRACT

The purpose of the current study was to investigate the effectiveness of four methods of handling missing data. Effectiveness was defined as the probability of reproducing the covariance matrix of the target sample accurately. Effectiveness of the missing data methods was assessed by manipulating the proportion of cases containing missing values and the sample size. The missing data methods studied were: (1) listwise deletion; (2) pairwise deletion; (3) regression; and (4) expectation maximization. The initial sample contained data from a study of 4,645 high school students who reported working while in school. This study replicated E. Witta's (2000) study with an addition to the sampling procedure. The pattern of missing values was used to create missing values in complete cases and compared to the case replacement method. Thus, the current study was also testing the effects of the case replacement sampling procedure used by Witta and the variable value replacement method used in this study. Results indicate there are differences in decisions concerning effectiveness of a missing data method due to the way the sample was created and the proportion of missing values used. (Contains 4 tables and 19 references.) (SLD)

Running Head: Missing Data Method

Does Method of Creating the Sample Influence

Missing Data Decisions?

E. Lea Witta

University of Central Florida

Department of Educational Foundations

Paper presented at the Annual Conference of the American Educational Research Association, Seattle, April 8-14, 2001.

Abstract

The purpose of the current study was to investigate the effectiveness of four methods of handling missing data. Effectiveness was defined as the probability of accurately reproducing the covariance matrix of the target sample. Effectiveness of the missing data methods was assessed by manipulating the proportion of cases containing missing values and the sample size. This study replicated Witta's (2000) study with an addition to the sampling procedure. The pattern of missing values was used to create missing values in complete cases and compared to the case replacement method. Thus, the current study was also testing the effects of the case replacement sampling procedure used by Witta and the variable value replacement method used in this study. Results indicate there are differences in decisions concerning effectiveness of a missing data method used due to the way the sample was created and the proportion of missing values used.

Does Method of Creating the Sample Influence

Missing Data Decisions?

When data are analyzed in survey research, often there are missing values. If the mechanism causing the missing values is known, the solution to this problem may be incorporated in the study. Inevitably, however, when data are collected by survey, subjects may fail to answer some questions for reasons unknown to the researcher. Ignoring this problem may lead to analysis of data that is of dubious value.

In addition, different methods of handling missing values may produce different results. When Jackson (1968) entered data on all the available variables in a discriminant analysis, the significance of the regression coefficients of individual variables, as well as the interpretation of the importance of these variables, changed with the missing value method used. Witta and Kaiser (1991) also reported that the regression coefficients and total variance accounted for by the variables changed depending on the method used to handle missing values. After re-analyzing three studies of private/public school achievement, Ward and Clark III (1991) concluded that the method used to handle missing data influenced the outcome of these studies.

When using the National Educational Longitudinal Study of 1988 database to investigate the effects of part-time work on school outcomes Singh and Ozturk (1999) eliminated more than half of the selected cases by listwise deletion of the incomplete data. The question then became was listwise deletion an appropriate method of for handling the missing data or, would another method be more effective? Witta (2000) investigated this question by stratifying the population into complete and incomplete cases and randomly selecting samples from each population to create data sets with varying proportions of incomplete cases. This method of sample selection suspect because the cases were then tested against the complete sample.

Statement of the Problem

The purpose of the current study was to determine if there was a difference in results if the incomplete cases were randomly selected and replaced complete ones or if values in complete cases were deleted (missing) based on the pattern of missing values in the incomplete sample. The effectiveness of four methods of handling missing data using the 26 variables in the Singh and Ozturk study was assessed under both sample creation conditions. Effectiveness was defined as the probability of accurately reproducing the true covariance matrix. Effectiveness of the missing data methods was assessed by manipulating the proportion of cases containing missing values and the sample size. The missing data methods studied were listwise deletion, pairwise deletion, regression and expectation maximization. Sample sizes investigated were 500, 1000, and 2000. The proportion of incomplete cases in each sample were 30%, 50%, and 70%.

Until recently, the only methods available with popular statistical computer software focused on handling the missing data problem by deleting subjects with incomplete information, deleting the variables with missing values, or replacing the missing value with some reasonable estimate. Now, however, new subroutines are available to provide more assistance in handling missing data and providing analysis choices using iterative regression or expectation maximization (EM) procedures. These relatively new methods (in current software) also provide the possibility of specifying the model to be used (i.e., multivariate normality, adding a randomly selected error).

Methods Studied

Listwise Deletion

Listwise deletion is probably the most frequently used method of handling missing data and is available as a default option in several statistical software programs including SPSS. This method discards cases with a missing value on any variable and thus is very wasteful of data. Listwise deletion, however, has been shown to be effective with low average intercorrelation, less than four variables and a small proportion of missing values (Chan, et.al., 1976; Haitovsky, 1968; Timm, 1970). The assumption of missing completely at random is crucial to the use of this method. It is more likely, however, to find the complete sample different in important ways from the incomplete sample (Little & Rubin, 1987). Problems for a researcher using this method include a reduction in power and an increase in standard error due to reduced sample size and the possible elimination of sub-populations.

Pairwise Deletion

When using pairwise deletion, covariances are computed between all pairs of variables having both observations, eliminating those that have a missing value for one of the two variables (Glasser, 1964). Means and variances are computed on all available observations. The assumption made is that the use of the maximum number of pairs and all the individual observations yield more valid estimates of the relationship between the variables. It is assumed that when two variables are correlated, information on one improves the estimates of the other variable. It is also assumed that the pairs are a random subset of the sample pairs. If these assumptions are true, pairwise deletion produces unbiased estimates of the variable means and variances (Hertel, 1976). When missing data are not missing completely at random, however, the correlation matrix produced by pairwise deletion may not be Gramian (Norusis, 1988).

Marsh (1998) investigated the estimates produced when using pairwise deletion for randomly missing data. From this study, which included five levels of missing data and three sample sizes, Marsh concluded parameter variability was explained, parameter estimates were unbiased, and only one covariance matrix was nonpositive definite.

Regression

Regression as an imputation method has many variations. The regression methods rely on information contained in non-missing values of other variables to provide estimates of missing values. In this procedure variables containing a missing value in one or more cases are regressed on the other variables. The resulting regression equation is used to provide an estimate which replaces the missing value. Then the process is repeated until estimates do not change.

As the average intercorrelation and the number of variables from which these methods can obtain information increases, the regression methods, theoretically, perform better. Too many variables, however, can cause problems with over prediction (Kaiser & Tracy, 1988) and too high an average intercorrelation can result in a singular matrix. In these cases, regression does not perform well.

Variations in the regression methods include differences in methods of developing the initial correlation matrix (listwise deletion, pairwise deletion, and mean substitution) and the presence or absence of iteration procedures. Differences in regression methods also include the use of randomly selected residuals for iterations and assumptions of a normal distribution. Theoretically, the more variables considered that provide additional information, the better the estimate. Mundfrom and Whitcomb (1998) investigated the effects of using mean substitution, hot-deck imputation, and regression imputation on classification of cardiac patients. Mean substitution and hot-deck imputation correctly classified patients more frequently than regression imputation.

Expectation Maximization

Dempster, Laird, and Rubin (1977) recommended the use of the EM algorithm which imputes estimates simultaneously in an iterative procedure. The E step of this algorithm finds the conditional expectation of the missing values. The M step performs maximum likelihood estimation as if there were no missing data. The primary difference between this procedure and the regression procedure is that the values for the missing data are not imputed and then iterated. The missing values are functions based on the conditional expectation (Little & Rubin, 1987). This method of handling missing data represents a fundamental shift in the way of thinking about missing data (Schafer & Olsen, 1998).

Pattern of Missing Values

All of the missing data handling procedures discussed require data missing at random (MAR) or missing completely at random (MCAR). Yet Cohen and Cohen (1983) suggested that in survey research the absence of data on one variable may be related to another variable and may be due to the value of the variable itself. When investigating simultaneously missing values, Witta (1996/97) found concurrently missing values ($p<.001$) in three of four samples using data from a national database.

Missing Completely at Random refers to no relationship between the missing value for one variable and missing values for other variables or between the missing value and values for other variables. The missing data subroutine in SPSS uses Little's test to evaluate this relationship. Because Missing at Random refers to the missingness due to the actual value of the variable (i.e., if salary is too high the respondent may refuse to answer), this procedure cannot be tested. Schafer and Olsen (1998), however, argue convincingly that "every missing-data method must make some largely untestable statistical assumptions about the manner in which the missing values were lost" (p551). Consequently, when analyzing real data, researchers typically assume missing at random.

Procedure

All high school seniors who had reported working during their senior year of high school and for whom base-year and first follow-up data were available were included in this study. The initial sample contained the 26 variables used in the Singh and Ozturk study for 4664 subjects. These subjects were split into three populations: those containing one or more missing values but less than 14 (n=1542), those containing more than 13 missing values (n=19), and those containing no missing values on any variable (n=3103). The 19 subjects having missing values for more than half the variables were eliminated from further analysis. The remaining two populations (n=4645) were used to create samples for analysis.

Creating Test Samples

A sample consisting of 2000 cases was randomly selected from the non-missing population. This sample was duplicated twice resulting in three identical samples of 2000 cases containing no missing values. The first sample was used to provide estimates of the target (true) covariance matrix.

A sample of 1400 cases was randomly select from the missing population. The pattern of missing values was recorded. This pattern was used to create missing values in 1400 of the complete cases in the second target sample. In addition, these cases were used to replace an equal number of randomly selected cases in the third target sample. This provided two test samples of 2000 with 70% of the cases containing missing values (one containing replacement cases and one having missing values created) and one target sample. This process was repeated to provide a test samples with 50% (1000) of the cases containing missing values. The process was repeated again to provide a test samples with 30% (600) of the cases containing missing values. This resulted in 9 samples of 2000 cases; 3 complete target samples and 2 test samples for each proportion of incomplete cases (30%, 50%, and 70%).

This entire procedure was repeated twice to provide test samples with 30%, 50%, and 70% of the cases containing missing values in test samples of 1000 and 500 cases. Thus, 27 samples were created; 9 test samples with replacement of the complete cases with incomplete ones, 9 test samples with missing values based on the pattern of the incomplete cases, and 9 complete target samples.

Analysis

Covariance matrices for the missing data handling methods were produced by the missing data subroutine in SPSS. The test for missing completely at random and pattern of missing data was also produced by this subroutine. After treatment by each missing data handling method, multi-sample analysis in LISREL 8.3 (Jöreskog & Sörbom, 1996, chap. 9) was used to test the equality of the covariance matrices produced by various missing data handling methods to the covariance matrix of the target sample.

Results

Randomness of Missing Values

Few of the samples used in the current study contained data missing completely at random. The frequency of simultaneously missing variables for each sample is presented in Table 1. The category of 'Test' consists of four simultaneously missing standardized test variables

(History, Math, Reading, and Science). The standardized test variables were also missing in conjunction with missing values for grades. The four grade variables were also missing simultaneously. If a variable did not contain a missing value for 10% of the sample cases (either alone or concurrently with other variables), it was included in the 'Other' category. In each sample, the majority of the cases containing missing values consisted of concurrently missing values for standardized tests (the categories 'Test' and 'Grades & Test').

Insert Table 1 About Here

Covariance Matrix Reproduction

All four missing data methods adequately reproduced (χ^2 p>.05) the target sample covariance matrix when 30% or 50% of the cases contained missing values regardless of sample size[1]. In addition, as depicted in Tables 2 and 3, the normed fit index (NFI) in all cases was above 0.98 and the root mean square residual (RMSEA) was 0.0. The listwise deletion and regression methods produced identical values for the two sampling conditions (replacement by pattern of missing values and replacement of cases by randomly selected cases). However, the pairwise deletion and EM algorithm methods produced much lower actual χ^2 values when the sample was created using replacement by pattern of missing values.

---

[1]To prevent discrepancies in sample size comparison, the n for testing the covariance matrices produced by Listwise and Pairwise deletion was enter in LISREL as the target n (i.e. if the target sample contained 500 cases, the n entered for the listwise deletion covariance matrix was 500).

Insert Tables 2 & 3 About Here

When complete cases were replaced with incomplete ones and 70% of the cases contained missing values, only the covariance matrix produced by the EM algorithm passably reproduced the target sample matrix when the sample size was 1000. Under the same conditions, when the sample size was 500 or 2000, no method adequately reproduced the target sample covariance matrix as measured by chi-square (χ^2, $p<.05$) . When, however, the pattern of missing values was used to create missing data in complete cases and 70% of the cases were incomplete, pairwise deletion and the EM algorithm adequately reproduced the target sample covariance matrix under all conditions. The normed fit index remained at an acceptable level of 0.96 or higher under all conditions. The root mean square residual also remained relatively small as shown in Table 4.

Insert Table 4 About Here

Discussion and Conclusions

When 30% or 50% of the cases in a sample were incomplete, all missing data methods tested adequately reproduced the target sample covariance matrix regardless of sample size. There was, however, a difference in the χ^2 statistic for the EM algorithm and pairwise deletion based on sampling method. In both instances if missing values were created in complete cases using the pattern of the incomplete ones, χ^2 was smaller. When 70% of the cases were incomplete only pairwise deletion and the EM algorithm could adequately reproduce the target sample covariance matrix when using the missing data pattern. This finding suggests that method of creating a sample for testing does indeed influence the results of a missing data analysis.

When using listwise deletion or regression, the method used to create missing values did not change results. This, of course, was due to the fact that listwise deletion uses only complete cases and in the comparative samples the complete cases were identical. Although the regression method used in this study was iterative, imputations were not used in each iteration and were based on the listwise deletion matrix. Consequently, the imputations were based upon the same data and were identical regardless of method used to create missing values.

Both pairwise deletion and the EM algorithm, however, used all known values. When the complete cases were replaced by incomplete ones, the values of the variables that were not missing for the incomplete case also replaced the variable values for the complete ones. When the missing data pattern was used to delete values from complete cases, the values that were not deleted did not change. Consequently, the pairwise deletion and EM algorithm estimates were based upon different values for each sample.

This study was limited to one sample size and proportion of incomplete cases for each test. Therefore, results may be specific to these samples. However, the results from this study show that replacement of complete cases with incomplete ones is an inappropriate method to use in study missing values. The results again indicate that if the proportion of incomplete cases is relatively small (30% or 50%) all methods used to handle missing values could adequately reproduce the covariance matrix of the target sample. Under these conditions, the researcher may choose which missing data handling method based upon substantive reasons. When the proportion of incomplete cases is large, only pairwise deletion and the EM algorithm could adequately reproduce the target covariance matrix. Under these conditions, the researcher should use one of these methods.

References

Chan, L.S., Gilman, J.A., & Dunn, O.J. (1976). Alternative approaches to missing values in discriminant analysis. *Journal of the American Statistical Association, 71*, 842-844.

Dempster, A.P., Laird, N.W., & Rubin, D.B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, B, 39*, 1-38.

Glasser, M. (1964). Linear regression analysis with missing observations among the independent variables. *Journal of the American Statistical Association, 59*, 834-844.

Haitovsky, Y. (1968). Missing data in regression analysis. *Journal of the Royal Statistical Society, B, 30*, 67-82.

Hertel, B.R. (1976). Minimizing error variance introduced by missing data in survey analysis. *Sociological Methods & Research, 4*, 459-474.

Hill, M.A. (1997). *SPSS Missing Value Analysis 7.5* [Computer program manual]. Chicago: SPSS Inc.

Jackson, E.C. (1968). Missing values in linear multiple discriminant analysis. *Biometrics, 24*, 835-844.

Jöreskog, K.G. & Sörbom, D. (1996). *Lisrel 8: User's reference guide*. Chicago: Scientific Software International, Inc.

Kaiser, J. & Tracy, D.B. (1988). Estimation of missing values by predicted score. *Proceedings of the Section on Survey Research, American Statistical Association 1988*. 631-635.

Little, R.J.A., & Rubin, D.R. (1987). *Statistical Analysis with Missing Data*. New York: John Wiley & Sons.

Marsh, H.W. (1998). Pairwise deletion for missing data in structural equation models: Nonpositive definite parameter estimates, goodness of fit, and adjusted sample sizes. *Structural Equation Modeling, 5* (1), p 22-36.

Mundfrom, D.J. & Whitcomb, A. (1998). *Imputing missing values: The effect on the accuracy of classification*. Paper presented at the annual meeting of the American Educational Research Association, San Diego. ED419817.

Norusis, M.J. (1988). *SPSS-X Introductory Statistics Guide: Release 3* [Computer program manual]. (pp 107-108). Chicago: SPSS Inc.

Schafer, J.L. & Olsen, M.K. (1998). Multiple imputation for multivariate missing-data problems: A data analyst's perspective. *Multivariate Behavior Research, 33* (4), p 545-571.

Singh K., & Ozturk, M. (1999). *Part-time work and school-related outcomes for high school seniors: An analysis of NELS:88*. Paper presented at the 1999 Annual Conference of the American Educational Research Association, Montreal, Canada.

Timm, N.H. (1970). The estimation of variance-covariance and correlation matrices from incomplete data. *Psychometrika, 35*, 417-437.

Ward, Jr., T.J. & Clark III, H.T. (1991). A reexamination of public-versus private-school achievement: the case for missing data. *Journal of Educational Research, 84*, 153-163.

Witta, E.L. (1996/97). Randomness of missing values in survey data. *Louisiana Education Research Journal, XXII* (2), p 73-86.

Witta L. & Kaiser, J. (1991, November). *Four methods of handling missing data with GSS-84*. Paper presented at the meeting of the Mid-South Educational Research Association, Lexington, KY

Tables

Table 1

Pattern of Missing Values by Proportion of Incomplete Cases and Sample Size

n	Grdes/Tests/40A-C	Hmwk10 School	Hmwk 10 Home	Hmwk 12 School	Hmwk 12 Schl/Hom	Absence 12	Grades 12	Grades & Tests	Tests 12	Other	χ^2	df
30%												
500[a]							20	34	54	42	880**	742
1000[b]							49	45	108	98	1581**	1230
2000[c]							99	118	199	184	1638.9*	1514
50%												
500[a]		7		6			33	40	92	72	1397**	1218
1000[b]		11					74	91	166	158	2402.1**	1998
2000[c]		33					156	160	348	303	2292.4	2235
70%												
500[a]	6	8		8	6	6	56	65	118	77	1280.1	1287
1000[b]		19				14	98	116	254	199	2260.8**	2017
2000[c]		42	21		35		209	242	515	357	2409.1	2414

Note. If less than 10% of a sample ([a]5, [b]10, [c]20) produced simultaneously missing values, the incomplete cases are recorded as other. χ^2 is Little's Missing Completely at Random. ** $p \leq .01$.

Table 2

Results when 30% of the Cases were Incomplete

Missing Data Method	Replacement - Pattern of Incomplete Cases			Replacement - Incomplete Cases		
	χ^2	RMSEA	NFI	χ^2	RMSEA	NFI
n=500						
Listwise	100.19	0.0	.99	100.19	0.0	.99
Pairwise	33.35	0.0	1.0	120.98	0.0	.99
EM	30.31	0.0	1.0	116.90	0.0	.99
Regression	99.66	0.0	.99	99.66	0.0	.99
n=1000						
Listwise	98.58	0.0	1.0	98.58	0.0	1.0
Pairwise	31.40	0.0	1.0	147.76	0.0	.99
EM	25.82	0.0	1.0	141.77	0.0	.99
Regression	99.37	0.0	1.0	99.37	0.0	1.0
n=2000						
Listwise	79.73	0.0	1.0	79.73	0.0	1.0
Pairwise	25.85	0.0	1.0	158.76	0.0	1.0
EM	23.03	0.0	1.0	150.99	0.0	1.0
Regression	81.91	0.0	1.0	81.91	0.0	1.0

Note. RMSEA = Root Mean Square Error of Approximation. NFI = Normed Fit Index. ** $p \leq .01$. * $p \leq .05$.

Table 3

Results when 50% of the Cases were Incomplete

Missing Data	Replacement - Pattern of Incomplete Cases			Replacement - Incomplete Cases		
Method	χ^2	RMSEA	NFI	χ^2	RMSEA	NFI
n=500						
Listwise	204.66	0.0	.98	204.66	0.0	.98
Pairwise	63.47	0.0	.99	230.41	0.0	.98
EM	52.24	0.0	1.0	209.12	0.0	.98
Regression	206.33	0.0	.98	206.33	0.0	.98
n=1000						
Listwise	194.10	0.0	.99	194.10	0.0	.99
Pairwise	54.18	0.0	1.0	233.52	0.0	.99
EM	47.26	0.0	1.0	224.22	0.0	.99
Regression	191.33	0.0	.99	191.33	0.0	.99
n=2000						
Listwise	205.20	0.0	.99	205.20	0.0	.99
Pairwise	54.50	0.0	1.0	326.51	0.0	.99
EM	43.90	0.0	1.0	313.80	0.0	.99
Regression	204.79	0.0	.99	204.76	0.0	.99

Note. RMSEA = Root Mean Square Error of Approximation. NFI = Normed Fit Index. ** $p \leq .01$. * $p \leq .05$.

Table 4

Results when 70% of the Cases were Incomplete

Missing Data Method	Replacement - Pattern of Incomplete Cases			Replacement - Incomplete Cases		
	χ^2	RMSEA	NFI	χ^2	RMSEA	NFI
n=500						
Listwise	515.79*	0.03	.96	515.79*	0.03	.96
Pairwise	114.87	0.0	.99	412.07*	0.02	.96
EM	104.86	0.0	.99	396.72*	0.01	.97
Regression	519.05*	0.03	.96	519.05*	0.03	.96
n=1000						
Listwise	444.59*	0.02	.98	444.59*	0.02	.98
Pairwise	116.84	0.0	1.0	402.53*	0.01	.98
EM	82.02	0.0	1.0	384.80	0.0	.98
Regression	441.73*	0.02	.98	441.73*	0.02	.98
n=2000						
Listwise	503.41*	0.02	.99	503.41*	0.02	.99
Pairwise	106.21	0.0	1.0	544.89*	0.02	.99
EM	92.26	0.0	1.0	529.91*	0.02	.99
Regression	510.22*	0.02	.99	510.22*	0.02	.99

Note. RMSEA = Root Mean Square Error of Approximation. NFI = Normed Fit Index. ** $p \leq .01$. * $p \leq .05$.

Table A-1

Study Questions and their suggested Construct

Construct	Variable Code	Question
Part-time Work	F1S85	HOW MANY HRS DOES R USUALLY WORK A WEEK
(Grades 10 & 12)	F2S88	CURRENT JOB, # HRS WORKED DURING SCHL YR
Attendance	F1S10A	HOW MANY TIMES WAS R LATE FOR SCHOOL
(Grade 10)	F1S10B	HOW MANY TIMES DID R CUT/SKIP CLASSES
	F1S13	HOW MANY DAYS WAS R ABSENT FROM SCHOOL
Attendance	F2S9A	HOW MANY TIMES WAS R LATE FOR SCHOOL
(Grade 12)	F2S9B	HOW MANY TIMES DID R CUT/SKIP CLASSES
	F2S9C	HOW MANY TIMES DID R MISS SCHOOL
Participation	F1S40A	OFTEN GO TO CLASS WITHOUT PENCIL/PAPER
(Grade 10)	F1S40B	OFTEN GO TO CLASS WITHOUT BOOKS
	F1S40C	OFTEN GO TO CLASS WITHOUT HOMEWORK DONE
Participation	F2S24A	GO TO CLASS WITHOUT PENCIL/PAPER
(Grade 12)	F2S24B	GO TO CLASS WITHOUT BOOKS
	F2S24C	GO TO CLASS WITHOUT HOMEWORK DONE
Homework	F1S36A1	TIME SPENT ON HOMEWORK IN SCHOOL
(Grade 10)	F1S36A2	TIME SPENT ON HOMEWORK OUT OF SCHOOL
Homework	F2S25F1	TOTAL TIME SPENT ON HMWRK IN SCHOOL
(Grade 12)	F2S25F2	TOTAL TIME SPENT ON HMWRK OUT SCHL
Grades 12	F2RHENG2	AVERAGE GRADE IN ENGLISH (HS+B)
	F2RHMAG2	AVERAGE GRADE IN MATHEMATICS (HS+B)
	F2RHSCG2	AVERAGE GRADE IN SCIENCE (HS+B)
	F2RHSOG2	AVERAGE GRADE IN SOCIAL STUDIES (HS+B)
Standardized Tests	F22XHSTD	HISTORY/CIT/GEOG STANDARDIZED SCORE
(Grade 12)	F22XMSTD	MATHEMATICS STANDARDIZED SCORE
	F22XRSTD	READING STANDARDIZED SCORE
	F22XSSTD	SCIENCE STANDARDIZED SCORE

Reproduction Release

(Specific Document)

TM032842

I. DOCUMENT IDENTIFICATION:

Title: Does Method of Creating the Sample Influence Missing Data Decisions	
Author(s): E. Lea Witta	
Corporate Source: American Educational Research Assoc. Conference	Publication Date: April 2001

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, Resources in Education (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign in the indicated space following.

The sample sticker shown below will be affixed to all Level 1 documents	The sample sticker shown below will be affixed to all Level 2A documents	The sample sticker shown below will be affixed to all Level 2B documents
PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY SAMPLE TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY SAMPLE TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)	PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY SAMPLE TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
Level 1	Level 2A	Level 2B
☑	☐	☐
Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g. electronic) *and* paper copy.	Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only	Check here for Level 2B release, permitting reproduction and dissemination in microfiche only

Documents will be processed as indicated provided reproduction quality permits.
If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche, or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Signature: E. Lea Witta	Printed Name/Position/Title: E. Lea Witta Associate Prof.	
Organization/Address: University of Central Florida PO Box 161250 Orlando, FL 32816-1250	Telephone: 407-823-3220	Fax: 407-823-5141
	E-mail Address: lwitta@mail.ucf.edu	Date: April 18, 2001

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
4483-A Forbes Boulevard
Lanham, Maryland 20706
Telephone: 301-552-4200
Toll Free: 800-799-3742
e-mail: ericfac@inet.ed.gov
WWW: http://ericfac.piccard.csc.com

EFF-088 (Rev. 9/97)