DOCUMENT RESUME

ED 312 597 CG 022 095

AUTHOR	Marsiske, Michael; Willis, Sherry L.
TITLE	Selective Attrition Effects in a Longitudinal Study of Adult Intelligence: Methodological Considerations.
SPONS AGENCY	National Inst. on Aging (DHHS/PHS), Bethesda, MD.
PUB DATE	18 Nov 89
GRANT	R01-AG05304
NOTE	37p.; Paper presented at the Annual Meeting of the Gerontological Society of America (42nd, Minneapolis, MN, November 17-21, 1989).
PUB TYPE	Reports - Research/Technical (143) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Adults; *Aging (Individuals); *Attrition (Research Studies); Intellectual Development; *Intelligence; Locus of Control; Longitudinal Studies; *Older Adults; *Research Problems; Statistical Bias

ABSTRACT

Selective subject attrition from longitudinal study panels can bias estimates of developmental change. Particularly in studies of older adults, sampling effects can adversely affect attempts to estimate true ontogenetic change. Selective attrition effects were examined in 636 Pennsylvania adults (138 males, 498 females), aged 58-91, who were te~ ed in 1978-1979; and 232 subjects who returned and were retested in 1986-1987. On both occasions, subjects received measures of intellectual ability, locus of control beliefs, and attitudes toward aging. Comparison of the Time One ability performances of returning and non-returning subjects indicated significantly lower performance levels for dropouts on measures of Verbal Ability, Figural Relations, Induction, Experiential Evaluation, Memory Span, and Perceptual Speed (p .05). Logistic multiple regression procedures identified significant control belief and demographic predictors of attrition status ("R"=.25) and attrition type ("R"=.327). The group of non-returning subjects tended to be older, contained a higher proportion of males, were more likely to believe that chance controlled intellectual performance, had lower intellectual achievement motivation levels, were less likely to be employed, and reported lower levels of subjective health. The effects of these predictors were partialled out of the abilities showing significant attrition effects; this eliminated the significant attrition effect on four of the six abilities. These results suggest a possible procedure for examining selective attrition effects and quantifying sample bias in longitudinal studies of the elderly. (Author/NB)

Selective Attrition Effects in a Longitudinal Study of Adult Intelligence: Methodological Considerations[1]

Michael Marsiske & Sherry L. Willis

Pennsylvania State University

S-113 Health and Human Development Building

Department of Human Development and Family Studies

University Park, PA 16802

Running Head: SELECTIVE ATTRITION: METHODOLOGY

1 Paper presented at the annual meeting of the 42nd Annual Scientific Meeting of the Gerontological Society of America, Minneapolis, Minnesota, November 18, 1989

This study was part of a research program entitled the Adult Development and Enrichment Project (ADEPT) funded by the National Institute on Aging (R01-AG05304)

Abstract

Selective attrition effects were examined in 636 Pennsylvania adults (138 males, 498 females), aged 58-91 (M = 70, SD = 6.25) tested in 1978-1979, and again in 1986-1987, when 232 subjects returned. At both occasion, subjects received measures of intellectual ability, locus of control beliefs, and attitudes toward aging. Comparison of the Time One ability performances of returning and non-returning subjects indicated significantly lower performance levels for dropouts on measures of Verbal Ability, Figural Relations, Induction, Experiential Evaluation, Memory Span, and Perceptual Speed (p < .05). Logistic multiple regression procedures identified significant control belief and demographic predictors of attrition status ("R" = .25) and attrition type ("R" = .327). Non-returning subjects tended to be older, contained a higher proportion of males, were more likely to believe that Chance controlled intellectual performance, had lower intellectual achievement motivation levels, were less likely to be employed, and reported lower levels of subjective health. The effects of these predictors were partialled out of the abilities showing significant attrition effects; this eliminated the significant attrition effect on four of the six abilities. These results suggest a possible procedure for examining selective attrition effects and quantifying sample bias in longitudinal studies of the elderly.

Selective Attrition Effects in a Longitudinal Study of Adult Intelligence: Methodological Considerations

Selective subject attrition from longitudinal study panels can bias estimates of developmental change, especially when there is correlation between repeated study participation and the measurement variables (Baltes, Schaie & Nardi, 1971). Particularly in studies of older adults, where mortality and the inability to continue would be expected to remove participants from an ongoing study, sampling effects can adversely affect attempts to estimate true ontogenetic change. When only survivors are studied in investigations of developmental change, the age-performance functions obtained may not represent those of the original parent population (Schaie, Labouvie & Barrett, 1973). Furthermore, external validity or representativity is not the only issue affected by selective attrition bias. Berk (1983) notes that selective sampling can also affect <u>internal validity</u>. He states, "By excluding some observations in a systematic manner, one has inadvertently introduced the need for an additional regressor [which explains the systematic biasing] (p. 388)". Models which fail to account for sample selectivity are therefore inherently misspecified.

Longitudinal studies examining health, demographic and lifestyle changes in aging have found that, in general, when returning subjects are compared to subjects who drop out, retested subjects tend to show more advantaged levels on a number of variables, including health status and Instrumental Activities of Daily Living (IADL) rankings, prescription drug use, social service use, incomes, life satisfaction levels, and mean ages (e.g., Horgas, Marsiske & Ahern, 1988; Markides, Dickson & Pappas, 1982; Wilson & Webber, 1976).

Numerous studies have also considered differences between returning and non-returning older adults in intellectual functioning. Jarvik & Falek (1963) found that the stability of intellectual performance on the WAIS was related to survival in a multi-wave study of older adults. Baltes, Schaie & Nardi (1971) and Schaie, Labouvie & Barrett (1973) found that returning subjects at a later occasion of measurement had higher mental ability levels, as well as greater cognitive and visuomotor flexibility, when compared to non-returning subjects at an earlier occasion of measurement. Researchers from the Duke Longitudinal Study reported, in a study of older adults that included 11 testing sessions over a 20-year period, that the successively smaller groups of returning subjects at each time of measurement had successively higher WAIS full scale scores (Siegler & Botwinick, 1979).

Cooney, Schaie & Willis (1988) reported that selective attrition effects on cognitive and personality variables were themselves selective. In their study, only 3 of 5 intellectual measures showed significant participation main effects: Verbal Meaning, Inductive Reasoning, and Psychomotor Speed. Selective attrition effects also varied by cohort; with the older of the two cohorts examined generally showing a stronger attrition effect. When examining Drop X Cohort interaction effects, participation effects were more evident on Crystallized abilities. These results suggested that attrition effects may not be universal across the abilities studied.

In another study related to intellectual development, Lachman & Leff (1989) found that survivors in a 5-year longitudinal study of perceived control and intellectual functioning in old age were younger and higher in verbal intelligence. Using the Personality-In-Contexts measure (Lachman,1986) the researchers also found that dropouts had stronger external control beliefs in the dimensions of Intellectual Chance and Powerful Others.

To this point, we have made the case that non-returning subjects in longitudinal studies of older adults tend to demonstrate lower functional levels than returning subjects. This conclusion proceeds from studies conducted across functional domains, and holds for health, demographic, attudinal and intellectual functioning. What has not been considered thus far are differences between dropout types. Some subjects leave an ongoing study for biological or "natural" reasons (e.g., illness, death), while others leave due to psychosocial or "unnatural" reasons (e.g., disinterest, moving away). There is not a homogenous "drop" group (Baltes, Schaie & Nardi, 1971; Goudy, 1985a).

In a study of Iowa elders, Powers & Bultena (1972) compared deceased subjects with subjects who returned at a follow-up 11 years later. Subjects who died had been older, and had lower incomes, poorer health, and were more likely to be unemployed prior to their deaths. Similarly, Riegel, Riegel & Meyer (1967) found, at a five-year follow-up of older German subjects, that persons who were less intellectually able were more likely to die or become seriously ill. This relationship was true for younger subjects only (55-64); the occurrence of death appeared to be more random at later age levels (65+). These findings were recently extended by White & Cunningham (1988). They compared deceased elderly Floridans who had failed to return for a longitudinal follow-up due to death with subjects who were matched in age and gender using a small psychometric ability battery. Under these cross-sectional comparison conditions, only Vocabulary showed an association with distance from death, and only for those who died at age 70 or earlier and who died within two years of testing. Thus, White & Cunningham reasoned, differences between living and deceased subjects appear to be neither universal (i.e., present across all variables investigated) nor dramatic.

A small number of studies have compared both kinds of drop groups (biological and psychosocial dropouts) with returning subjects. In the non-intellectual domain, Norris (1985) found that healthiest subjects endured longest in a longitudinal study of Kentucky adults aged 55 and over. Norris also found that subjects who failed to return due to death, disability or an untraceable change of address showed disadvantages relative to returning subjects in mental and physical health and social support, but the dropouts who quit because of disinterest did not differ. Horgas, Marsiske & Ahern (1988) reported similar findings. In the early wave of their study, both survey non-respondents and deceased subjects were in poorer health, and used more drugs and services than returnees. Subjects who died, however, tended to be sicker and tended to use more drugs than other non-respondents. When comparing the different kinds of dropout Markides, Dickson & Pappas (1982) found that refusers were the most advantaged of the dropouts. Subjects who could not be located showed few differences from deceased subjects, and those who were too ill to participate were even worse off than the deceased.

In the intellectual domain, Cooney, Schaie & Willis (1988) provided one of the few systematic comparisons of attrition types. Their results suggested that attrition effects were largest for subjects who departed for biological reasons. Attrition effects were also greater for older subjects, leading the authors to suggest that young voluntary (i.e., psychosocial) departers from a longitudinal study may be relatively less biasing to the ontogenetic change trajectories obtained.

The issues of the nature and magnitude of biases introduced by selective subject attrition in longitudinal studies merit further consideration here. To begin with, no matter how discrepant a non-returning subgroup is from a returning subgroup, it is unlikely to introduce bias if the proportion of non-returning subjects is small (Markides, Dickson, & Pappas, 1982). One of the lowest attrition rates for a study of adults is reported by Rose, Bosse & Szretter (1976). Their study of disease-free men across the adult lifespan included younger adult men, who would be expected to have lower attrition rates. Over the 12-year interval reported in their study, only 11% of the initial population was lost. This yields a mean annual attrition rate of 0.92%. Selectivity at the outset of a study may minimize natural subject selection processes, but the sample will remain selected nonetheless; here, the results are generalizable only to a disease-free population of men.

Streib (1966), Goudy (1985b) and Norris (1987) all examined differences in health, attitude, and demographic variable intercorrelations between returning and non-returning subgroups. These analyses were premised on the notion that if variable relationships were the same in both subgroups at an initial occasion, then selective attrition could not be said to have exerted a strong differentiating influence on the groups. Supporting such a notion, despite strong participation effects, there was little difference in variable intercorrelations for the different participation groups. It is important to note, however, that invariant patterns of covariation across time or across groups do not preclude the possibility of level changes in the observed variables, or in the latent variables they represent (Gorsuch, 1983). Furthermore, finding structural similarity in a set of variables between returning and non-returning subject groups at an initial time of measurement does not address whether variable intercorrelations would be similar at a later time of measurement for either subgroup. Thus, findings that variable relationships do not differ between returning and non-returning subject groups says nothing about group similarities or differences in ontogenetic change trajectories, nor does it address the central issue of whether selective subject attrition introduces bias into the results obtained with returning subjects only.

The present study attempts to address a number of the issues surrounding sample bias and selective subject attrition. First, we will examine the nature and rate of attrition in the ADEPT (Adult Development and Enrichment Project at Penn State) longitudinal study over a 7-year interval. Second, we will examine the universality of attrition effects over the intellectual measures examined. Third, we will attempt to model the attrition effect (both at the crude drop/return level, and also by examining differences between biological and psychosocial dropouts), to determine what control beliefs and personal factors help to account for selective subject loss. Fourth, we will examine the effects of controlling for these factors on differences between returning and non-returning subgroups. Finally, we will attempt to assess how controlling for selective attrition biases may affect the longitudinal change trajectories we observe.

Method

Sample

The initial sample, first tested in 1978-1979, was composed of 636 community-dwelling Caucasian Pennsylvania adults (138 males, 498 females). The subjects ranged in age from 58 to 91 years (M = 70.0 years, SD = 6.25 years). Self-reported years of education ranged from 2 to 22 years (M = 11.45 years, SD = 3.02 years). Most of the sample reported itself as unemployed, retired, or homemakers (92.7%); 7.3% of subjects reported that they were still working full- or part-time. 38.9% of the sample was married; the remaining 61.1% were widowed, single, divorced or separated.

Subjects were asked to rate their vision, hearing and general physical health. The scale ranged from 1 (very good) to 6 (very poor). The mean vision self-rating was 2.35 (SD = 0.95). Mean self-reported hearing was 2.36 (SD = 1.11). Self-reported health was reported at a mean level of 2.04 (SD = 0.89).

At a 1986-1987 follow-up, 232 subjects returned to participation. The reasons for non-return were as follows: 116 subjects did not return for psychosocial reasons (Moved away, not located, family illness, not interested), 135 subjects did not return due to illness and infirmity (vision problems, too ill, in nursing home), and 133 subjects did not return due to death. An additional 20 subjects were not eligible to return, due to failure to attend all testing sessions in 1978-1979, sensory deficits, or below-chance performance levels in 1978-1979. These subjects were excluded from all further analyses. Thus, the total attrition from the study constituted 63.5% of the sample, and the annual attrition rate was 9.1%.

Each subject was paid approximately $2.00 per hour of participation at the initial testing sessions; payment was made either directly to the participant or the community centre in which the subject was tested, according to the subject's wishes.

Subjects who returned in 1986-1987 received the same test battery and the same testing/training procedures that they had received in 1979. Subject payments had been increased to $5.00 per hour.

Procedure

Subjects were assessed on the ADEPT Ability Battery (see below). Tests were administered to small groups of subjects (n=4-13). A tester and a proctor supervised each testing session. The ADEPT Study was comprised of a number of sub-studies; although all subjects received essentially the same tests, different sub-studies received different testing orders and different test forms. The analyses in the present paper report only pretest data (some subjects received interventions and posttests as well, but these data are not included here).

Measures

Ability battery. The ability battery was developed within the fluid (Gf) and crystallized (Gc) model of intelligence (cf. Cattell, 1971), and multiple marker tests of seven primary dimensions of intelligence were included. These first-order dimensions included Induction, Figural Relations, Verbal Ability, Experiential Evaluation, Semantic Relations, Perceptual Speed, and Memory Span.

Table 1 shows the marker tests for each ability. Fluid intelligence (Gf) was represented by the primary abilities of Induction (I) and Figural Relations (CFR). These marker tests required subjects to find relationships within arrays of letters and numbers, and figures, respectively. Crystallized intelligence (Gc) was represented by the primary abilities of Verbal Ability (V), and Experiential Evaluation (EMS). The Verbal measures were recognition vocabulary tests; the EMS tests required subjects to solve problems of a social nature. One marker ability loaded on both Gf and Gc; marker tests of this ability (Semantic Relations (CMR)) required subjects to complete verbal analogies. Perceptual Speed (Ps) measures assessed the speed with which subjects made simple visual discriminations; Memory Span (Ms) measures assessed subjects' digit span, or working memory.

Insert Table 1 about here

With the exception of two measures, all tests used were adapted versions of previously published psychometric ability tests. The adaptations included enlargement of test stimuli, simplification of response format, and reduction of the number of items. Reduction of test length was based on the internal consistency of the full-length measures; all reduced test versions had reliabilities greater than .65. Two measures were developed specifically for the ADEPT Study; these were the ADEPT Induction Test and the ADEPT Figural Relations Diagnostic Test (Blieszner, Willis & Baltes, 1981; Willis, Blieszner & Baltes, 1981).

For the analyses in this paper, the best marker (i.e. the measure with the highest loading on the primary ability it represents, as identified by Baltes et al (1980) and longitudinally confirmed by Willis & Jay (1989)) was selected to represent an individual's performance on that factor; since different sub-studies of the ADEPT study received different testing protocols, the measure used to represent each primary ability may vary for sub-studies. For this reason, analyses reported in this paper refer to the primary ability examined, and not the specific measure used.

Intellectual control and self-efficacy beliefs measure. Intellectual control was assessed by the Personality-In-Contexts (PIC) Inventory. The PIC was developed for the ADEPT study, to measure older adults' beliefs and attributions regarding their own cognitive functioning (Lachman, 1986). The PIC has six scales. Three of these scales are Locus of Control scales (Internal, Chance, and Powerful Others); these scales measure older adults' evaluations of their control of intellectual competence. In addition, a PIC Anxiety scale assesses the presence of anxiety associated with intellectually demanding tasks. The PIC Attitude toward Own Intellectual Aging Scale assesses perceived changes in intellectual competence that are associated with aging. The PIC Achievement scale assesses the perceived importance of intellectual accomplishments.

Personal data questionnaire. Subjects provided information about age, marital status, education, physical health, vision and hearing ability.

Derivation of scores

Standardization of scores. Since different subgroups of subjects received different Forms of the same measures, all measures were standardized to a mean of 50, and a standard deviation of 10. This permitted different Forms of the same test to be compared using a common metric; it also permitted cross-test comparisons. Since the standardization procedure used total group means as the base, subgroup differences (e.g., between returning and non-returning subjects) were preserved. Second occasion (1986-1987) data for returning subjects was also standardized; each measure was scaled using 1978-1979 scores for the total sample as a base. This facilitated comparison of the magnitude of age-related change across measures. Standardization was performed on both psychometric ability measures and the PIC.

Results

Differences between returning and non-returning subjects. To examine differences between returning and non-returning subjects in ability test performance, independent group t-tests were conducted on the marker test scores of the 7 primary abilities. There were significant differences between returning and non-returning subjects on six of the seven abilities, with returning subjects having higher scores than non-returning subjects: Verbal ability ($t(544) = -3.27$, $p < .01$); Figural Relations ($t(625) = -6.78$, $p < .0001$); Induction ($t(621) = -4.02$, $p < .001$); Experiential Evaluation ($t(257) = -3.28$, $p < .01$); Memory Span ($t(258) = -2.59$, $p < .01$); and Perceptual Speed ($t(543) = -4.70$, $p < .0001$). Only the Semantic Relations measure did not show a significant difference between returning and non-returning subjects.

Differences between returning subjects and attrition sub-groups. Unbalanced one-factor analyses of variance were conducted on the 7 ability marker tests. Drop Status (returned, died, too ill to participate, psychosocial reason for non-participation) was a between-subjects factor. Main effects of Drop Status were found on 5 of the 7 abilities: Verbal ability ($F(3,529) = 7.03$, $p < .0001$); Figural Relations ($F(3,610) = 16.61$, $p < .0001$); Induction ($F(3,606) = 5.38$, $p < .01$); Experiential Evaluation ($F(3,247) = 4.41$, $p < .01$); and Perceptual Speed ($F\ (3,528) = 9.19$, $p < .0001$). No main effect of Drop Status was found for Semantic Relations or Memory Span.

For the variables showing significant effects of Drop Status, Tukey's "honestly significant difference tests" were conducted on cell means, to discern differences between attrition subgroups. Due to the unbalanced design of the ANOVA, comparisons were done using Tukey-Kramer's adjustment for unequal cell sizes. All significant mean differences reported below are significant at $p < .05$. Variable levels of the Drop Status variable will be referred to as "Returning" (subjects who returned to the study), "Ill" (subjects who did not return to the study due to illness), "Dead" (deceased subjects), and "Psycho-social" (subjects who did not return to the study for non-biological reasons).

For Verbal ability, Ill subjects had significantly lower scores than Returning subjects and Psycho-social dropouts. Dead subjects did not differ from any of the other subject groups. On Figural Relations, all three attrition types showed significantly lower scores than returning subjects, but none of the three attrition types were significantly different from one another. For Inductive Reasoning, Experiential Evaluation, and Perceptual Speed, both the Dead and the Ill showed significantly lower performance levels than Returning subjects, but psychosocial dropouts did not differ from any of the other attrition groups. Table 2 shows the cell means and standard deviations for each Drop Status group.

Insert Table 2 about here

Prediction of attrition. In order to determine which variables were related to attrition status (i.e., whether or not a subject returned to the study in 1986-1987), regression analyses were conducted. Since the Drop variable was dichotomous (0=not returned, 1=returned), and the Drop Status variable was polytomous (0=Return, 1=Psychosocial attrition, 2=Illness, 3=Death), Ordinary Least Squares regression was not appropriate (Hanushek & Jackson, 1977). Multiple logistic regression analyses examined the relationship between demographic and control belief variables to the probability of remaining in the study. Although ability variables might have accounted for a substantial proportion of the variance in attrition (since most showed significant attrition effects), psychometric ability variables were not included in any of the models predicting attrition, so that non-ability predictors of attrition could later be partialled out of ability scores (in an attempt to control for attrition effects on ability data). To avoid the loss of many cases due to listwise deletion (i.e., loss of subjects who have one or more missing values on any variable entered into model), missing value dummies (0=missing value, 1=no missing value) were coded for each variable (Clogg, Note 1).

For the model regressing the binary Drop variable on personal and control belief predictors, significant parameter estimates were obtained for PIC Achievement (r = .13), PIC Intellectual Chance (r = -.08), Age (r = -.09), Employment status (working/not working) (r = .10), and self-rated Health (r = -.06). These correlations suggest that persons with lower Achievement Motivation, higher Chance orientation, greater age, unemployment and worse health tend to drop out.

For the model predicting the polytomous Drop Status variable, significant parameter estimates were obtained for Age (r = .10), Gender (1=male, 2=female) (r = -.09), Health (r = .09), PIC Achievement (r = - .09), Employment status (working/not working) (r = -.05), and PIC Intellectual Chance (r = .05). Although the direction of effects appears to be opposite to that obtained above, this is because the polytomous Drop Status variable was coded in the opposite direction of the binary Drop variable.

The results of these analyses are presented in Table 3. For both models, parameter estimate convergence was achieved in five iterations, with no step halvings. Both models were significant; for the dichotomous attrition variable (Drop), $X^2(5) = 43.38$, $p < .0001$, for the polytomous attrition variable (Drop Status) $X^2(6) = 73.72$, $p < .0001$. While an actual R^2 estimate is not possible for logistic regression, since the homoscedasticity assumption does not hold, a "pseudo R^2" (Aldrich & Nelson, 1984) was computed for each model. For both models, explained variance was low. The model predicting the binary Drop variable had a pseudo R^2 of .07; the model predicting the polytomous Drop Status variable had a pseudo R^2 of .11.

Insert Table 3 about here.

Controlling for attrition predictors in ability variables. The demographic and control belief predictors of attrition and attrition type identified above were entered into Ordinary Least Squares regressions, in which each of the ability variables showing significant attrition effects was regressed on the predictors of attrition, thereby attempting to control for the attrition effects.

For the dichotomous (Drop) attrition variable, t-tests on standardized (*M* = 50, *SD* = 10) residualized ability variables (i.e., ability variables with the demographic and control belief predictors of attrition partialled out) were conducted, to compare returning and non-returning subjects. Group differences were no longer significant for two of the six ability variables previously showing significant attrition effects (Verbal ability and Memory Span). The remaining ability variables (Figural Relations, Induction, Experiential Evaluation, and Perceptual Speed) continued to show significant differences between returning and non-returning subjects ($p < .05$). Table 4 displays the means and standard deviations for the six ability variables showing significant attrition effects, both before and after the predictors of attrition were partialled out.

Insert Table 4 about here

For the polytomous (Drop Status) attrition variable, unbalanced ANOVAs and Tukey's comparison of means were conducted on the five variables that showed a main effect of Drop Status. Two variables no longer showed a significant effect of Drop Status (Experiential Evaluation, Induction). While the main effect of Drop Status remained significant ($p < .05$) on the other three variables, there were differences in the post hoc comparisons of means for two of these abilities: For Figural Relations, both the Ill and the Dead remained significantly lower than Returning subjects, but Psycho-social dropouts no longer had significantly lower scores than returning subjects. Similarly, for Perceptual Speed, only the Ill subjects remained significantly lower than Returning subjects; Dead subjects had residualized scores not significantly different from Returning subjects. Table 2 above shows the cell means for each of the ability variables showing attrition effects, prior to residualization. The standardized means and standard deviations of the ability variables, after the demographic and control belief predictors of attrition were partialled out, are presented in Table 5.

Insert Table 5 about here

Effects of residualization on the estimation of developmental change. For the four ability variables on which attrition differences were reduced by partialling out predictors of attrition, we sought to examine whether controlling for selective attrition altered estimates of developmental change. Another set of regression analyses was performed; the predictors of attrition which had reduced significant differences between returning and non-returning subjects on Verbal Ability, Memory, Induction, and Experiential Evaluation were simultaneously partialled out of both the Time One and Time Two ability scores for returning subjects. The resulting residual distributions were transformed into T-scores (M = 50, SD = 10), using the means and standard deviations of the Time One residual distributions as the standardization base. Table 6 presents the mean 1978-1979 and 1986-1987 ability scores, both before and after controlling for identified Time One attrition predictors. The magnitude of mean ability decline estimated from residualized scores is not appreciably different from that obtained with the original scores.

Insert Table 6 about here

Discussion

As with many other studies of selective subject attrition in longitudinal research with older adults, the results of the present study suggest that non-returning subjects are generally worse off than returning subjects intellectually, demographically, and in terms of control beliefs. Significant attrition effects were found on six of the seven psychometric ability variables examined in the present study. In all cases, non-returning subjects had significantly lower scores than returning subjects. Returning subjects also had higher PIC Achievement Motivation Scores, and a lower likelihood of attributing their intellectual performance to Chance (not unlike the results of Lachman & Leff, 1989). Returning subjects were also more likely to be younger, employed, and female. Returning subjects also reported higher subjective health levels.

The present study also provides some support for the notion of non-homogeneity of drop groups. While, generally, non-returning subjects performed at lower levels than returning subjects, <u>post hoc</u> comparisons of the different drop groups suggested that the dead and ill dropouts differed from returning subjects more than psycho-social dropouts did. Further, in those comparisons where dead subjects differed from ill dropouts, we found that the Ill subjects were actually <u>more</u> impaired than the Dead. This is similar to findings by Norris (1985). Like Cooney, et al (1988) and Norris (1985), we also found that persons who left the study for non-biological reasons tended not to differ significantly from returning subjects, supporting the notion that non-biological dropouts tend to be less biasing to a longitudinal study. Thus, while Goudy (1985a) notes that non-biological dropouts are <u>theoretically</u> more biasing to a longitudinal study (because, unlike dead subjects who leave both a study <u>and</u> the parent population, non-biological dropouts leave the study only), in practice psycho-social dropouts appear to be less biasing.

Our attempts to model the attrition effect met with limited success. In part due to our decision not to include psychometric ability variables as predictors in our models (since we wanted to use non-intellectual predictors of attrition as controls for the attrition effect in the ability variables), we accounted for relatively little variance in the attrition dummy variables, even though the models accounted for significantly greater than zero variance. This is an important finding, however, regarding the usefulness of our approach. Clearly, success in modelling the attrition effect is largely a function of the variables selected for inclusion in a study. The better the data base on subject background and personal characteristics, the better the prediction of attrition is likely to be. Unfortunately, there are prices to pay for including many personal variables. First, the variables may be tangential to the main research questions, and may therefore increase the expense and time of data collection. Second, as the number of variables increases relative to sample size, the likelihood of finding significant effects by chance alone increases. Thus, from a practical perspective, it may never be possible to account for a large proportion of the variance in the attrition effect without including ability variables.

Interestingly, despite the relatively low variance in attrition that we were able to account for, we were still able to reduce differences between returning and non-returning subjects in all ability variables, simply by controlling for the predictors of attrition we had identified. We actually eliminated the significance of the difference between returning and non-returning subjects in two ability variables (Verbal Ability and Memory Span). We also eliminated the significance of the main effect of Attrition Type in two other variables (Inductive Reasoning and Experiential Evaluation). These results suggest that we were successful in identifying some of the predictors of differences between returning and non-returning subjects, in at least 4 of the ability variables under study. Interestingly, these results also suggest that different predictors are differentially salient in accounting for differences between returning and non-returning subjects on different ability variables; since different personal variables may account for subjects' performance on any one ability measure, it follows that different personal variables may also account for attrition group differences on any one ability variable.

Partialling identified predictors of attrition out of both 1978-1979 and 1986-1987 ability variables produced interesting results. While controlling for certain variables did reduce differences between returning and non-returning subjects, it had no effect on our estimates of developmental change. That is, the magnitude of age related decline on residualized ability variables was not appreciably different from that obtained with the original standardized ability variables. This is encouraging, because it suggests that the variables which help to account for ability differences between returning and non-returning subjects at Time One exert only a small effect on the ability performance of returning subjects at Time Two. This would appear to suggest that longitudinal outcomes for returning subjects are not largely biased by attrition group differences at Time One. This conclusion needs to be tempered by the fact that we were not able to account for a large proportion of the variance in attrition; it could be argued that if we had identified more non-ability predictors of the attrition effect, these variables might also have exerted a larger effect on Time Two performance levels for returning subjects. More importantly, these results really do not address the central questions: What would the developmental change trajectory look like if there were no dropouts? How would the dropouts have changed if we could have observed them (e.g., if we had had more measurement occasions).

In this sense, our results support Rogosa's (1988) assertion that "attempts to statistically adjust for preexisting differences...[are] doomed to failure (p. 190)".

Although the results reported in this paper do not directly enable one to examine the effect of selective attrition bias on estimates of ontogenetic change, they provide further support that one can identify the nature of the bias and quantify it. They also suggest that non-biological dropouts from a longitudinal study are less biasing than subjects who drop out for biological reasons. This raises an interesting theoretical issue: If, practically speaking, non-biological dropouts are not very biasing to the representativity of a sample, and if biological subject mortality in a longitudinal panel is <u>representative of mortality occurring within the population</u> (Goudy,1985a), does it follow that differences between returning and non-returning subjects do <u>not</u> dramatically bias the estimation of ontogenetic change? (Note that if lowest-functioning members of a population tend not to volunteer for even the initial phases of a longitudinal study, then population mortality may be underestimated by sample mortality.) The present results suggest that returning subjects <u>are</u> biased representatives of the larger population they represent, largely because of <u>ill</u> dropouts in a longitudinal study. Ill dropouts have the lowest intellectual levels of all drop groups. Norris' (1985) results also suggest that ill dropouts are the most impaired. <u>Ill dropouts leave a longitudinal study, but not always the broader population</u>. Thus, for both empirical and theoretical reasons, it appears that the loss of these most impaired <u>ill</u> subjects is most biasing to longitudinal studies, particularly studies of older adults, where the proportion of ill subjects can be high.

Reference Notes

1. Clogg, C. C. Personal communication.

Missing data values typically cause the loss of a substantial number of cases in regression modelling, due to listwise deletion. Further, if different subjects have missing values on each predictor, sample size may vary with the model being considered, making it difficult to compare models.

The approach being considered here codes a dummy variable, D, for <u>each variable entered into the model</u>, with a value of 1 for each subject with no missing values, and 0 for each subject with a missing value. Missing values are not recoded into System-specific missing codes (e.g., "." in SAS, "M" in AIDA), but left as originally coded (e.g., "99").

Consider the simple case of one dependent variable , Y, and two predictors, X_2, X_3. We create two dummy variables D_2, and D_3. $D_2 = 0$ if a subject has a missing value on X_2 , and $D_2 = 1$ if X_2 is not missing. $D_3 = 0$ if a subject has a missing value on X_3 , and $D_3 = 1$ if X_3 is not missing. Now, instead of simply regressing Y on X_2 and X_3, we perform the following regression:

$$Y = b_1 + b_2 X_2 D_2 + b_3 X_3 D_3$$

To illustrate how this approach avoids listwise deletion, consider the following two cases: I) Subject has no missing values, II) Subject has a missing value (99) for X_2.

Case I) Both dummy variables equal 1 (for no missing values); therefore, they add nothing to the regression equation. The D terms drop out, and :

$$Y = b_1 + b_2 X_2 + b_3 X_3$$

Case II) Dummy D_2 equals 0 (zero), since the subject has a missing value for X_2. Thus, D_2 multiplied by X_2 causes the X_2 term to drop out, adding nothing to the value of Y. D_3 still equals 1, since X_3 is not missing, and adds nothing to the regression equation, causing the D_3 term to drop out as well. IMPORTANTLY, this subject is <u>not</u> lost from the analysis simply because he/she does not have a value for a variable. The regression equation for this subject is:

$$Y = b_1 + b_3 X_3$$

For the subject in this second case, the estimated value of Y is based on the <u>available</u> information, .

Note also that the model can be expanded to determine whether there are significant differences between subjects with and without missing values; this provides a practical test of whether subjects with missing values should in fact be pooled with subjects who have no missing values. Here, the dummy variables themselves are included in the regression model:

$$\underline{Y} = b_1 + b_2 \underline{D}_2 + b_3 \underline{X}_2 \underline{D}_2 + b_4 \underline{D}_3 + b_5 \underline{X}_3 \underline{D}_3$$

Here, if parameter estimates of b_2 or b_4 are significantly greater than zero, this indicates that subjects with missing values on either variable are significantly different from subjects without missing values on either variable. Note also that pooling subjects with and without missing values assumes that the two groups are drawn from the same population, and that the variables common to both groups have homogenous variances.

References

Aldrich, J. H., & Nelson, F. D. (1984). Linear probability, Logit, and Probit Models. Newbury Park, CA: Sage.

Baltes, P. B., Cornelius, S. W., Spiro, A., Nesselroade, J. R., & Willis, S. L. (1980). Integration versus differentiation of fluid/crystallized intelligence in old age. Developmental Psychology, 16, 625-635.

Baltes, P. B., Schaie, K. W., & Nardi, A. H. (1971). Age and mortality in a seven-year longitudinal study of cognitive behavior. Developmental Psychology, 5, 18-26.

Berk, R. A. (1983). An introduction to sample selection bias in sociological data. American Sociological Review, 48, 386-398.

Blieszner, R., Willis, S. L., & Baltes, P. B. (1981). Training research in aging on the fluid ability of inductive reasoning. Journal of Applied Developmental Psychology, 2, 247-265.

Cattell, R. B. (1971). Abilities: Their structure, growth, and action. Boston: Houghton-Mifflin.

Cattell, R. B., & Cattell, A. K. S. (1957). Test of "g": Culture Fair (Scale 2,Form A). Champaign, IL: Institute for Personality and Ability Testing.

Cattell, R. B., & Cattell, A. K. S. (1961). Measuring intelligence with the Culture-Fair Tests: Manual for Scales 2 and 3. Champaign, IL: Institute for Personality and Ability Testing.

Cattell, R. B., & Cattell, A. K. S. (1963). Test of "g": Culture Fair. (Scale 3, Form A and Form B). Champaign, IL: Institute for Personality and Ability Testing.

Cooney, T. M., Schaie, K. W., & Willis, S. L. (1988). The relationship between prior functioning on cognitive and personality dimensions and subject attrition in longitudinal research. Journal of Gerontology: Psychological Sciences, 43, P12-P17.

Ekstrom, R. B., French, J. W., Harman, H., & Derman, D. (1976) Kit of factor-referenced cognitive tests (Revised ed.). Princeton, NJ: Educational Testing Service.

Gorsuch, R. L. (1983). Factor Analysis (Second Edition). Hillsdale, NJ: Erlbaum.

Goudy, W. J. (1985a). Effects of sample attrition and data analysis in the Retirement History Study. Experimental Aging Research, 11, 161-167.

Goudy, W. J. (1985b). Sample attrition and multivariate analysis in the Retirement History Study. Journal of Gerontology, 40, 358-367.

Guilford, J. P. (1969a). Verbal Analogies Test, I. Beverly Hills, CA: Sheridan Psychological Services.

Guilford, J. P. (1969b). Word Matrix Test. Beverly Hills, CA: Sheridan Psychological Services.

Hanushek, E. A., & Jackson, J. E. (1977). Statistical Methods for Social Scientists. Orlando, FL: Academic Press.

Horgas, A. L., Marsiske, M., & Ahern, F. (1988). Attrition in a large scale longitudinal study of the elderly: The role of health-status variables. Paper presented at the 41st Annual Scientific Meetings of the Gerontological Society of America, San Francisco, CA.

Horn, J. L. (1967). Social Situations - EP03A. Unpublished test, University of Denver, 1967.

Jarvik, L. F., & Falek, A. (1963). Intellectual stability and survival in the aged. Journal of Gerontology, 18, 173-176.

Lachman, M. E. (1986). Personal control in later life: Stability, change and cognitive correlates. In M. M. Baltes & P. B. Baltes (Eds.) The Psychology of Control and Aging (pp. 207-236). Hillsdale, NJ: Erlbaum.

Lachman, M. E., & Leff, R. (1989). Perceived control and intellectual functioning in the elderly: A 5-year longitudinal study, Developmental Psychology, 25, 722-728.

Markides, K. S., Dickson, H. D., & Pappas, C. (1982). Characteristics of dropouts in longitudinal research on aging: A study of Mexican Americans and Anglos. Experimental Aging Research, 8, 163-167.

Norris, F. H. (1985). Characteristics of older nonrespondents over five waves of a panel study. Journal of Gerontology, 40, 627-636.

Norris, F. H. (1987). Effects of attrition on relationships between variables in surveys of older adults. Journal of Gerontology, 42, 597-605.

O'Sullivan, M., & Guilford, J. P. (1965). Social Translations, Form A. Beverly Hills, CA: Sheridan Psychological Services.

O'Sullivan, M., Guilford, J. P., & de Mille, R. (1965) Measurement of Social Intelligence (Technical Report Number 34). Los Angeles: University of Southern California, Psychology Laboratory.

Plemons, J. K., Willis, S. L., & Baltes, P. B. (1978). Modifiability of fluid intelligence in aging: A short-term longitudinal training approach. Journal of Gerontology, 33, 224-231.

Powers, E. A., & Bultena, G. L. (1972). Characteristics of deceased dropouts in longitudinal research. Journal of Gerontology, 27, 530-535.

Raven, J. C. (1962). Advanced Progressive Matrices, Set II (revised). London: H. K. Lewis.

Riegel, K. F., Riegel, R. M., & Meyer, G. (1967). A study of the dropout rates in longitudinal research on aging and the prediction of death. Journal of Personality and Social Psychology, 5, 342-348.

Rogosa, D. (1988). Myths about longitudinal research. In K. W. Schaie, R. T. Campbell, W. Meredith, & S. C. Rawlins (Eds.) Methodological Issues in Aging Research (pp. 171-210). New York: Springer.

Rose, C. L., Bosse, R., & Szretter, W. T. (1976). The relationship of scientific objectives to population selection and attrition in longitudinal studies: The case of the Normative Aging Study. Gerontologist, 16, 508-516.

Schaie, K. W., Labouvie, G. V., & Barrett, T. J. (1973). Selective attrition effects in a fourteen-year study of adult intelligence. Journal of Gerontology, 28, 328-334.

Siegler, I. C., & Botwinick, J. (1979). A long-term longitudinal study of intellectual ability in older adults: The matter of selective subject attrition. Journal of Gerontology, 34, 242-245.

Streib, G. F. (1966). Participants and drop-outs in a longitudinal study. Journal of Gerontology, 21, 200-209.

Thurstone, T. G. (1962). Primary Mental Abilities for Grades 9-12 (Revised edition). Chicago: Science Research Associates.

White, N., & Cunningham, W. R. (1988). Is terminal drop pervasive or specific? Journal of Gerontology: Psychological Sciences, 43, P141-P144

Willis, S. L., Blieszner, R., & Baltes, P. B. (1981). Intellectual training research in aging: Modification of performance on the fluid ability of figural relations. Journal of Educational Psychology, 73, 41-50.

Willis, S. L., & Jay, G. M. (1989). Structual invariance of cognitive abilities from young-old to old-old age. Unpublished manuscript, The Pennsylvania State University, University Park, PA.

Wilson, A. J. E., & Webber, I. L. (1976). Attrition in a longitudinal study of an aged population. Experimental Aging Research, 2, 367-387.

Table 1.

ADEPT Measurement battery: Hypothesized general intelligence dimensions, primary abilities, and marker tests

General dimension	Primary ability	Test	Source
Gf	CFR	Culture-Fair Test (Scale 2 Form A) and Power Matrices (Scale 3, Form A, 1963 ed., and Form B, 1961 ed.)	Cattell & Cattell (1957, 1961, 1963)
	CFR	ADEPT Figural Relations Diagnostic Test (Form A)	Plemons, Willis & Baltes (1978)
	CFR	Raven's Advanced Progressive Matrices (Set II)	Raven (1962)
Gf	I	ADEPT Induction Diagnostic Test (Form A)	Blieszner, Willis & Baltes (1981)
	I	Induction Standard Test	Ekstrom, French, Harman, & Derman (1976); Thurstone (1962)
Gf/Gc	CMR	Verbal Analogies I	Guilford (1969a)
	CMR	Word Matrix	Guilford (1969b)
Gc	EMS	Social Translations (Form A)	O'Sullivan & Guilford (1965); O'Sullivan, Guilford & deMille (1965)
	EMS	Social Situations (EP03A)	Horn (1967)
Gc	V	Verbal Meaning (9-12)	Thurstone (1962)
	V	Vocabulary (V-2, V-3, V-4)	Ekstrom et al (1976)
Ms	Ms	Visual Number Span	Ekstrom et al (1976)
	Ms	Auditory Number Span	after Ekstrom et al (1976)
	Ms	Auditory Number Span-Delayed Recall	after Ekstrom et al (1976)
Ps	Ps	Finding A's	Ekstrom et al (1976)
	Ps	Number Comparisons	Ekstrom et al (1976)
	Ps	Identical Pictures	Ekstrom et al (1976)

Table 2.

Mean performance of each attrition sub-group on the ability variables showing significant attrition main effects.

Ability	Mean: (SD):	Returning[a]	Psycho-Social[b]	Ill[c]	Dead[d]
Verbal Ability		51.71 (9.56)	51.65 (9.45)	46.99 (9.79)	49.00 (10.26)
Figural Relations		54.21 (9.16)	50.72 (9.30)	47.66 (10.19)	48.62 (10.20)
Inductive Reasoning		52.06 (9.55)	49.73 (8.28)	48.60 (6.99)	49.64 (8.56)
Experiential Evaluation		52.48 (10.43)	50.82 (10.43)	48.00 (8.55)	47.31 (9.12)
Perceptual Speed		51.98 (9.61)	50.16 (10.83)	47.26 (8.24)	47.16 (9.62)

[a] Returning = Subjects who returned to the study in 1986-1987

[b] Psycho-social = Subjects who left the study for non-biological reasons

[c] Ill = Subjects who were too ill to participate in 1986-1987

[d] Dead = Subjects who were deceased at 1986-1987 follow-up

Table 3.

Summary of multiple logistic regression analyses: PIC control belief and personal variables as predictors of attrition

Dependent variable	Significant predictors	b-weights	X^2	Model R
Drop (0=dropout, 1=return)	PIC Achievement	0.029	16.63***	
	PIC Chance	-0.019	6.66**	
	Age	-0.028	8.93**	
	Employment Status	0.397	10.97***	
	Health	-0.208	5.04*	0.254
Drop Status (0=return 1=psycho-social drop reason 2=illness 3=death)	Age	0.042	19.63***	
	Gender	-0.660	13.88***	
	Health	0.310	15.12***	
	PIC Achievement	-0.025	15.83***	
	Employment Status	-0.239	5.81*	
	PIC Chance	0.015	6.24*	0.327

Table 4.

Initial and residualized scores for returning and non-returning subjects on ability variables showing attrition effects.

Ability		Original Mean	Original S.D.	Residualized Mean	Residualized S.D.
Verbal Ability	Returning	51.71	9.56	50.88	9.38
	Non-ret.	48.89	9.99	49.43	10.36
Figural Relations	Returning	54.21	9.16	52.47	9.09
	Non-ret.	48.79	9.94	48.56	10.23
Inductive Reasoning	Returning	52.06	9.55	51.18	10.85
	Non-ret.	49.22	7.88	49.30	9.41
Experiential Evaluation	Returning	52.48	10.43	51.48	10.71
	Non-ret.	48.46	9.21	48.94	9.35
Memory Span	Returning	53.66	8.50	50.85	8.80
	Non-ret.	50.64	9.75	49.40	10.76
Perceptual Speed	Returning	51.98	9.61	51.71	10.09
	Non-ret.	48.01	9.60	48.89	9.79

Table 5.

Mean performance of each attrition sub-group on t ability variables showing significant attrition main effects, after partial g out of attrition predictors.

Ability	Mean: (SD):	Returning[a]	Psycho-Social[b]	Ill[c]	Dead[d]
Verbal Ability		50.93 (9.31)	51.53 (9.67)	47.42 (10.67)	50.28 (10.30)
Figural Relations		52.52 (9.06)	49.82 (9.63)	47.53 (10.68)	48.75 (10.34)
Inductive Reasoning		51.22 (10.84)	49.07 (9.71)	48.74 (8.58)	50.38 (10.10)
Experiential Evaluation		51.44 (10.71)	49.68 (10.01)	48.43 (9.76)	49.04 (8.96)
Perceptual Speed		51.64 (10.06)	50.01 (11.17)	48.17 (8.49)	49.04 (9.95)

[a] Returning = Subjects who returned to the study in 1986-1987

[b] Psycho-social = Subjects who left the study for non-biological reasons

[c] Ill = Subjects who were too ill to participate in 1986-1987

[d] Dead = Subjects who were deceased at 1986-1987 follow-up

Table 6.

Longitudinal change on four ability variables, before and after partialling out predictors of attrition.

		Original Mean	Original S.D.	T	Residual Mean	Residual S.D.	T
Verbal Ability	Time One	51.71	9.56	-2.84** (df=213)	51.93	9.64	-2.87** (df=213)
	Time Two	50.50	9.88		50.68	9.99	
Memory Span	Time One	53.75	8.58	1.69 (df=101)	51.15	9.18	1.68 (df=101)
	Time Two	55.60	11.54		53.17	12.56	
Inductive Reasoning	Time One	52.06	9.54	-3.81*** (df=230)	52.13	10.93	-3.91*** (df=230)
	Time Two	50.61	8.25		50.39	9.47	
Experiential Evaluation	Time One	52.32	9.88	-3.86*** (df=101)	51.89	9.89	-3.89*** (df=101)
	Time Two	48.67	8.79		48.17	8.61	

** $p < .01$
*** $p < .001$