SPATIAL TRAINING IN LATER ADULTHOOD: COMPONENTS OF CHANGE

Sherry L. Willis and K. Warner Schaie

The Pennsylvania State University

The authors gratefully acknowledge the assistance of Ann O'Hanlon wit
data analyses. Paper presented at the American Educational Research
Association, Chicago, April, 1985.

SPATIAL TRAINING IN LATER ADULTHOOD:

COMPONENTS OF CHANGE

Introduction

For the past decade my colleagues and I have been involved in a program of research examining the modifiability of cognitive performance in later adulthood (Baltes & Willis, 1982; Schaie & Willis, 1978; Willis, 1985). It has been our position that descriptive research on cognitive development, which has focused on normative patterns of change, should be supplemented by experimental studies examining the plasticity of cognitive performance in later adulthood (Willis & Baltes, 1980). Our training research has focused on those abstract reasoning abilities, which within the classical pattern of cognitive aging, are said to show relative early normative decline (Botwinick, 1977).

What are some of the major conclusions from our prior training research? First, older adults exhibited significant training improvement for a variety of reasoning abilities studied (figural relations, inductive reasoning, attention processes). Second, training improvement appears to be ability-specific. That is, the elderly's cognitive performance was improved on several measures of the ability focused on in training (e.g., inductive reasoning); however, there were no significant training effects for measures of other abilities (e.g., vocabulary) that were not focused on in training. Third, maintenance of training effects was demonstrated over a six-month period.

While the findings from these early training studies were encouraging, a number of important questions remain to be addressed. One critical question focuses on whether training is effective in remediating prior decline in older adults' performance or whether training improves the performance level of those older adults who have not shown significant age-related decline. In contrast to training research with children which assumes that the performance levels acquired through training had not previously been demonstrated, it is not possible to determine the nature of training effects in old age, unless there exists prior longitudinal data to define the previous cognitive performance of the older adult.

Descriptive longitudinal research indicates that while normative patterns of significant intellectual decline begin in the mid sixties, there are a number of older adults who do not exhibit decline until the seventies and a few do not decline until the eighties (Schaie, 1983). Given the cross-sectional design of prior cognitive training research, it has not been possible to address the question of the extent to which training is effective in remediating decline vs improving the performance of those who have not declined. There has been the implicit assumption that training remediates decline, but this assumption thus far has not been tested. To our knowledge the

study we are presenting offers the first available data on the issue of the relationship between decline status and training effects.

A second important question focuses on what specific components of performance change as a function of training. Several alternative hypotheses regarding the nature of training improvement in old age can be specified. One hypothesis is based on the finding of a general slowing of response speed with age (Birren, Woods, & Williams, 1980); training improvement is hypothesized to reflect primarily an increase in number of attempted items, but no change in accuracy. A second hypothesis is that training of cognitive strategies does result in improved accuracy, as reflected in an increase in correct responses and a decrease in commision errors.

This paper presents data from a cognitive training study conducted with older adults to address these two issues. The study involved training on two cognitive abilities, Inductive Reasoning and Spatial Orientation. We will report findings only for Spatial Orientation training.

Description of training study

Subjects.

During the last two and one-half years we have been involved in conducting training research with older participants from the Seattle Longitudinal Study, a cohort-sequential longitudinal study that has examined patterns of intellectual change in adulthood over the past 28 years (Schaie, 1983). Subjects in our training reseach were 229 older adults (M = 97; F = 132) over the age of 60 who had been participants in the longitudinal study over the prior 14 years (1970-1984) or longer. Mean age of the sample was 72.8 years (Range = 64-95); mean educational level was 13.9 years (Range = 6-20). All subjects were community dwelling. Prior to the initiation of the study, each subject's physician was contacted and asked to indicate whether the subject suffered any known physical or mental disabilities which would interfere with participation in the study; subjects so identified were not included in the study.

Definition of decline status.

The decline status of subjects was determined by the following procedure. Subject's test performance on the Thurstone PMA Reasoning

remained
1970-1984).
one
nts; Space
r interval
luded the
fects would
hird

and Spatial Orientation measures were classified as having stable or having declined over the prior 14 year interval. The statistical criterion for the definition of decline was standard error of measurement or greater (Reasoning = 4 po = 6 points). Determination of decline status over a 14 ye involving three data points (1970, 1977, 1984 pretest) pre confound of regression toward the mean, since regression e not continue in the same direction from the second to the

measurement points (cf. Nesselroade, Stiegler & Baltes, 1980). In 1970, prior to the onset of decline, there was no significant difference between stable and decline subjects on Space; the decline subjects actually performed significantly better ($p < .02$) than the stable subjects on the Reasoning measure in 1970.

Forty-six percent (N = 107) of the sample were classified as having remained stable on both ability measures; fifteen percent (N = 35) of the subjects had declined on Reasoning, but not on Space; sixteen percent (N = 37) had declined on Space but not on Reasoning; and approximately twenty-two percent (21.8%; N = 50) had declined on both measures (Figure 1). As would be expected, stable subjects (M = 70.9 years) were somewhat younger than decline subjects (M = 74.4 years; $p < .001$). Decline and stable subjects did not differ significantly on educational level or income.

Insert Figure 1 about here

Subjects were assigned to either Reasoning or Space training programs, based on their performance status. Subjects who had declined on Reasoning, but not on Space, or vice versa were assigned to the training program for the ability exhibiting decline. Subjects who had remained stable on both abilities or had shown decline on both abilities were randomly assigned to one of the training programs. Space training subjects included 51 (M = 23; F = 28) stables and 67 decliners (M = 29; F = 38). Reasoning training subjects included 56 (M = 25; F = 31) stables and 55 (M = 20; F = 35) decliners. Stable subjects in Space training were significantly younger ($p < .$
[illegible]
better educated ($p < .05$) than decliners. Stable subjects in Reasoning training were also significantly younger ($p < .001$) than decliners, but decliners were significantly better educated ($p < .01$) than stables.

Design of the study.

The study involved a pretest-treatment- posttest control group design. The Reasoning training group served as a treatment control for the Space training group and vice versa. All subjects had previously participated in the ISLE, and were informed via a series of letters that a new phase of the study was beginning. Subjects indicating an interest in participation were visited in their homes by a staff member. The purpose of the home visit was to discuss details of the study and to answer questions, to assess sensory handicaps which might interfere with participation, and to determine whether the home was a suitable place for conducting the training sessions. Subjects were administered a broad psychometric ability battery in 2 pretest sessions (2 1/2 hrs per session). Based on their prior longitudinal performance plus their pretest scores, subjects were assigned to either the Reasoning or Space training program.

Training involved 5 one-hour individually conducted training sessions. The majority of subjects were trained in their homes. Two middle-aged trainers, with prior educational experience in working with adults, served as trainers. Subjects were randomly assigned to the trainers within pragmatic constraints, such that each trainer trained approximately equal numbers of stable and decline subjects in each training program. Following training, subjects were assessed on a posttest battery involving the same measures administered at pretest. Subjects were paid $100 for participation in the study.

Measures.

A pre-posttest battery of psychometric measures representing five primary mental abilities was administered. In this paper we will present data only for the PMA measure of Spatial Orientation, for which longitudinal data are available. The PMA Space test is a multiple [illegible] The subject is shown a model line drawing and asked to identify whi of 6 choices shows the model drawn in different spatial orientation There are 2 or 3 correct responses possible for each test item. exemplar item is shown in Figure 2. Five scores per subject were derived: Correct Responses (Number of correct responses), Correct Nonresponse (Subject correctly did not mark a mirror image), Wrongs (Number of incorrect responses), Omits (Number of correct choices which should have been marked and were not, within the pool of item attempted by the subject), and Attempts (Total number of items attemted). There are 20 items on the PMA test with 6 answer choices per item. The corrects, unmarked correct responses, wrongs, and om sum to a total possible of 120 responses.

Insert Figure 2 about here

Training programs

The focus of the training was on facilitating the subject's us of effective cognitive strategies identified in previous research o the respective abilities.

A content task analysis of the PMA Spatial Orientation test wa conducted to identify the angle of rotation for each answer choice (Figure 3). Practice problems were developed to represent the angl rotations identified in the task analysis (45, 90, 135, 180 degrees Cognitive strategies to facilitate mental rotation which were focus upon in training included: 1) Development of concrete terms for various angles; 2) Practice with manual rotation of figures prior t mental rotation; 3) Practice with rotation of drawings of concrete, familiar objects prior to introduction of abstract figures; 4) Subject-generated names for abstract figures; and 5) Having the

subject focus on two or more features of the figure during rotation. These cognitive strategies had been identified in prior descriptive research on mental rotation ability (Cooper & Shepard, 1973; Egan, 1981; Kail, Pellegrino, & Carter, 1980).

Insert Figure 3 about here

Analyses.

To examine training effects, a 2 Training (Reasoning, Space) x 2 Status x 2 Gender x 2 Occasion ANOVA with repeated measures was performed.

For PMA Space, there were significant main effects for Status ($p < .001$), Gender ($p < .02$), and Occasion ($p < .001$). There were significant interactions for Training x Occasion ($p < .004$) and Training x Status x Occasion. The Training x Occasion interaction indicated a significantly higher performance for the Space training group at posttest. The triple interaction with Status reflects greater training gain for the decliners at posttest. Post hoc tests on PMA Space gain scores indicated that were significantly greater ($p < .01$) gains for decliners than for stables, for women than for men ($p < .01$), for women decliners than for women stables ($p < .02$), and for women decliners than for male decliners ($p < .03$).

Results

The above analyses indicate that there were significant training effects for subjects trained on Space, when compared with a control group. Our discussion of the training results will focus on differential training effects for two groups: Stables vs decliners, and men vs women We will report the results first for total scores and then for component scores.

Decline Status

Training improvement: Total scores. Our data indicate that training is particularly effective in improving the performance of decliners, as reflected in the interaction effects and post hoc analyses of the ANOVA reported above. Figure 4 presents the pre-posttest gain scores (in T-score points) for stable and decline subjects trained on each ability. The differential training effects in favor of decline subjects is most notable for Space training. On Space training 55% of the decline subjects exhibited significant pre-posttest training gain, compared with 39% of the stable subjects.

Insert Figure 4 about here

We have been able to examine further the practical significant of training effects for decline subjects by determining the extent to which training remediated these subjects to a prior performance level. That is, what proportion of the declining

in rights attributable to change in other components. About 69% of the increase in rights is attributable to an increase in the number of items attempted. Approximately 20% of the increase in rights is attributable to a decrease in omissions, and approximately 12% of the increase in rights is due to a decrease in wrong responses.

Figure 7 about here

If we compare the data in this way for Stable vs Decline subjects, we find different patterns in what score components account for training improvement (Figure 8). For stable subjects, 95% of the increase in correct responses is due to an increase in items attempted. In contrast, for Decline subjects, only 48% of the increase in correct responses is due to an increase in items attempted. For Decliners a decrease in omits accounts for 31% of the increase in rights, and a decrease in wrong responses accounts for 21% of the increase in rights. Thus, for Stable subjects the training improvement in correct responses is primarily associated with an increase in the items attempted, while for Decline subjects training improvement reflects not only an increase in number of items attempted, but also a decrease in wrong responses and omitted responses.

Insert Figure 8

Net change in component scores. The above analyses focusing on the increase in the number of correct responses are confounded, however, due to the significant increase in the number of items attempted from pre- to posttest. To examine net change in accuracy, we estimated the mean expected values for correct responses, no responses, wrongs, and omits at posttest based on performance at pretest. The proportions of attempted items at pretest attributed to correct response, no-responses, wrongs, and omits were computed. These proportions were then applied to obtain the expected values corrects, no-responses, wrongs, and omits at posttest. The differences between these expected values and the obtained values posttest were then computed. These values represent net changes i accuracy above and beyond those changes which would be expected to occur with the increase in attempted items. Figure 9 shows the ne change in rights, no-responses, wrongs, and omits for Stable and Decline subjects. For Stables, training improvement is reflected an increase in rights and a decrease in omits. For Decliners net change in accuracy involves both an increase in correct responses a decrease in wrongs.

Insert Figure 9 about here

Gender Differences

Training Improvement: Total score. There is a gender difference in the magnitude of training gain in the Space training, as indicated by the post hoc analyses of the ANOVA. Figure 10 (right side) presents gain scores, measured in T-score points, for men and women trained on Space. On Space training, women gained significantly more than did men.

Insert Figure 10 about here

Training Improvement: Component Scores Men's and women's performance on the five component scores was compared. Figure 11 shows for men and women the pre-posttest raw score gain in the number of correct, no-response, wrongs, omit, and attempted items. There were significant gender differences in gain in rights ($p < .02$) and no-responses ($p < .02$), and a trend ($p < .08$) toward a difference in attempted items. All of these gender differences were in favor of women. There was no gender difference in the drop in wrong responses. Thus, gender differences in training improvement reflect an increase in correct responses, but no differences in commission errors.

Insert Figure 11 about here

Proportion of change in rights attributable to change in other components. The proportion of pre-posttest change in each component type was computed in order to examine the relative contribution of the various score components to the increase in correct responses. The procedure for examining the proportion of change in rights attributable to change in other component scores was described above. Figure 12 presents these proportions. For men approximately 60% of the increase in rights is attributable to an increase in number of items attempted, and about 30% of the increase is attibutable to a decrease in omits. For women, approximately 69% of the increase in rights is attributable to an increase in the number of items attempted, while only 15% of the increase in rights is attributable to a decrease in omits. The training improvement (increase in rights) attributable to a decrease in wrong responses is roughly equivalent for the men and women. Thus, for both men and women the proportion change in rights is primarily due to an increase in the number of items attempted. However, the major gender difference is in the role of omits. For men, the decrease in the number of omits accounts for twice as much of the proportional increase in rights as is found for women (Men - 30%, Women - 15%).

Insert Figure 12

Net change in component scores. The final question deals with gender differences in the net change in accuracy, when the increase in the number of items attempted is controlled. The procedure for computing net accuracy change was described above. Figure 13 shows the net change in rights, no-responses, wrongs, and omits for men and women. The most salient gender difference was net change in wrong responses, with women showing a greater decrease in wrong responses than did men. If accuracy is defined as an increase in rights and a decrease in wrongs, then the net change in accuracy was greater for women than for men.

Insert Figure 13 about here

Summary

The primary purpose of this paper was to examine the nature of performance improvement associated with Space training for a sample of community-dwelling elderly. We examined training performance improvement with regard to two individual difference variables: decline status and gender differences.

Elderly subjects were classified as having remained stable or having declined in performance over the previous 14-year interval on a spatial task of mental rotation. Both decline and stable subjects received training on spatial orientation. There was a significant training effect; training was particularly effective for decline subjects. The nature of training improvement was examined for stable and decline subjects with regard to five component scores (rights, no-response, wrongs, omits, attempts). For decline subjects training resulted in an increase in accuracy of performance, as reflected by both a decrease in wrong responses and an increase in correct responses. Thus, training improvement for decliners was not solely a function of an increase in the number of items attempted (the speed hypothesis), but also reflected increased accuracy. However, for stable subjects training improvement reflected primarily an increase in correct responses, but no significant change in commission errors.

This difference in the nature of training improvement for stables versus decliners is reflected in the pie charts showing the change in proportion of right responses attributable to changes in the other score components. Ninety-five percent of the proportional increase in rights for stable subjects was attibutable to the increase in attempted items. In contrast, about half of the change in rights for the decliners was attributable to decreases in omits and commission errors. Finally, when the increase in the number of items attempted at posttest is controlled, it is again evident that training resulted in both an increase in correct responses and a decrease in commission errors for the decliners, while net changes for the stables were reflected in rights and omits scores.

Gender differences in the nature of training improvement were also examined. Women exhibited significantly greater pre-posttest gains in correct responses and in attempted items, while there was no significant gender differerence in decreases in wrong responses. However, when changes in net accuracy are examined, controlling for increases in the number of items attempted at posttest, it becomes evident that there is a greater improvement in accuracy for women than for men, as assessed by both an increase in rights and a decrease in wrongs.

These analyses suggest that there are individual differences in the nature of training improvement, and training effects need to examined with regard to these variables. Future research needs to examine the interaction between these individual difference variables and various training procedures.

References

Baltes, P. B. & Willis, S. L. (1982). Enhancement (plasticity) of intellectual functioning in old age: Penn State's Adult Development & Enrichment Project (ADEPT). In F. I. M. Craik & S. E. Trehub (Eds.), *Aging and cognitive processes.* New York: Plenum Press.

Cooper, L. B. & Shepard, R. N. (1973). Chronometric studies of rotation of mental images. In W. G. Chase (Ed.), *Visual information processing.* New York: Academic Press.

Egan, D. (1981). An analysis of spatial orientation test performance. *Intelligence, 5,* 85-100.

Holzman, T. G., Pellegrino, J. W., & Glaser, R. (1982). Cognitive dimensions of numerical rule induction. *Journal of Educational Psychology, 74,* 360-373.

Kail, R., Pellegrino, J., & Carter, P. (1980). Developmental changes in mental rotation. *Journal of Experimental Child Psychology, 39,* 102-116.

Kotovsky, K., & Simon, H. A. (1973). Empirical tests of a theory of human acquisition of concepts for sequential patterns. *Cognitive Psychology, 4,* 399-424.

Nesselroade, J. R., Stigler, S. M., & Baltes, P. B. (1980). Regression toward the mean and the study of change. *Psychological Bulletin, 88,* 622-637.

Schaie, K. W. (1983). The Seattle Longitudinal Study: A 21-year exploration of psychometric intelligence in adulthood. In K. W. Schaie (Ed.), *Longitudinal studies of adult psychological development* (pp. 64-135). New York: Guilford.

Schaie, K. W. & Willis, S. L. (1978). Life-span development: Implications for education. In L. Shulman (Ed.), *Review of research in education* (Vol VI). Itasca, IL: F. E. Peacock.

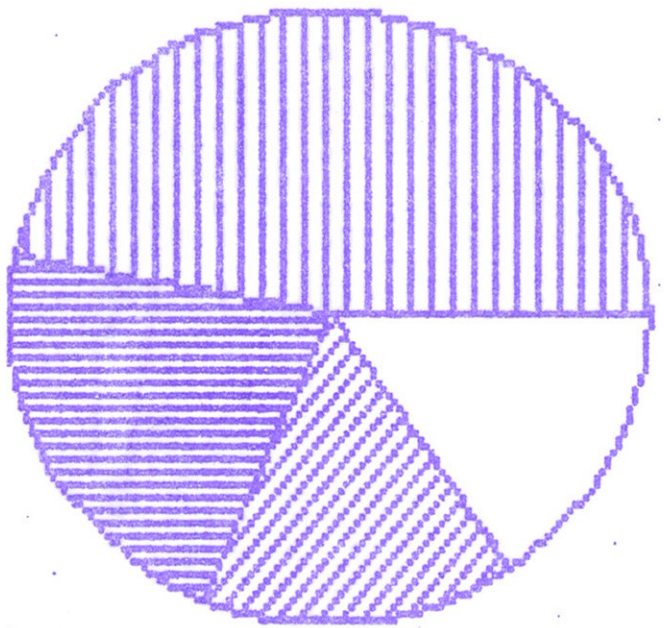
Thurstone, L. L. (1948). *Primary mental abilities.* Chicago: University of Chicago Press.

Willis, S. L. (1985). Towards an educational psychology of the adul learner. In J. E. Birren & K. W. Schaie (Eds.), *Handbook of th psychology of aging,* 2nd ed. New York: Van Nostrand Reinhold.

Willis, S. L. & Baltes, P. B. (1980). Intelligence in adulthood and aging: Contemporary issues. In L. Poon (Ed.), *Aging in the 1980s: Psychologicl issues.* Washington, DC: American Psychological Association.

STATUS: TRAINING SUBJECTS

DECLINE FROM 1970 TO 1984



15.30%
REASON. DECL.

16.20%
SPACE DECLINE

21.80%
BOTH DECLINE

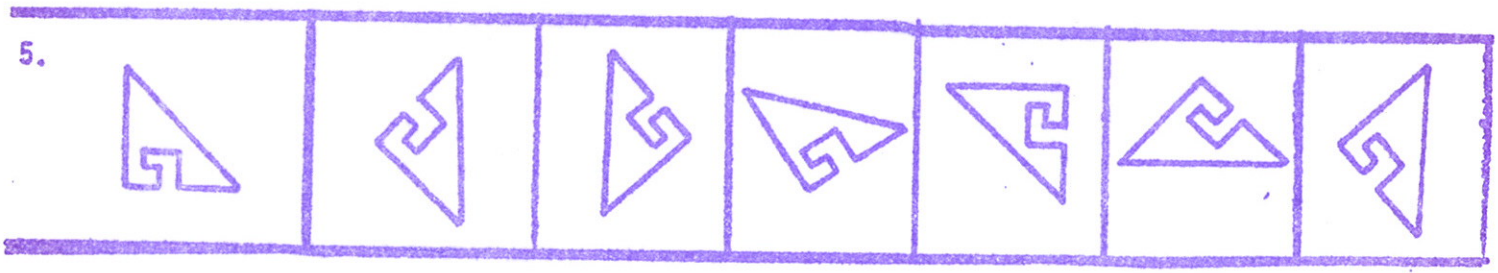
46.70%
STABLE

P M A REASONING

a z b y c - - - d w x e

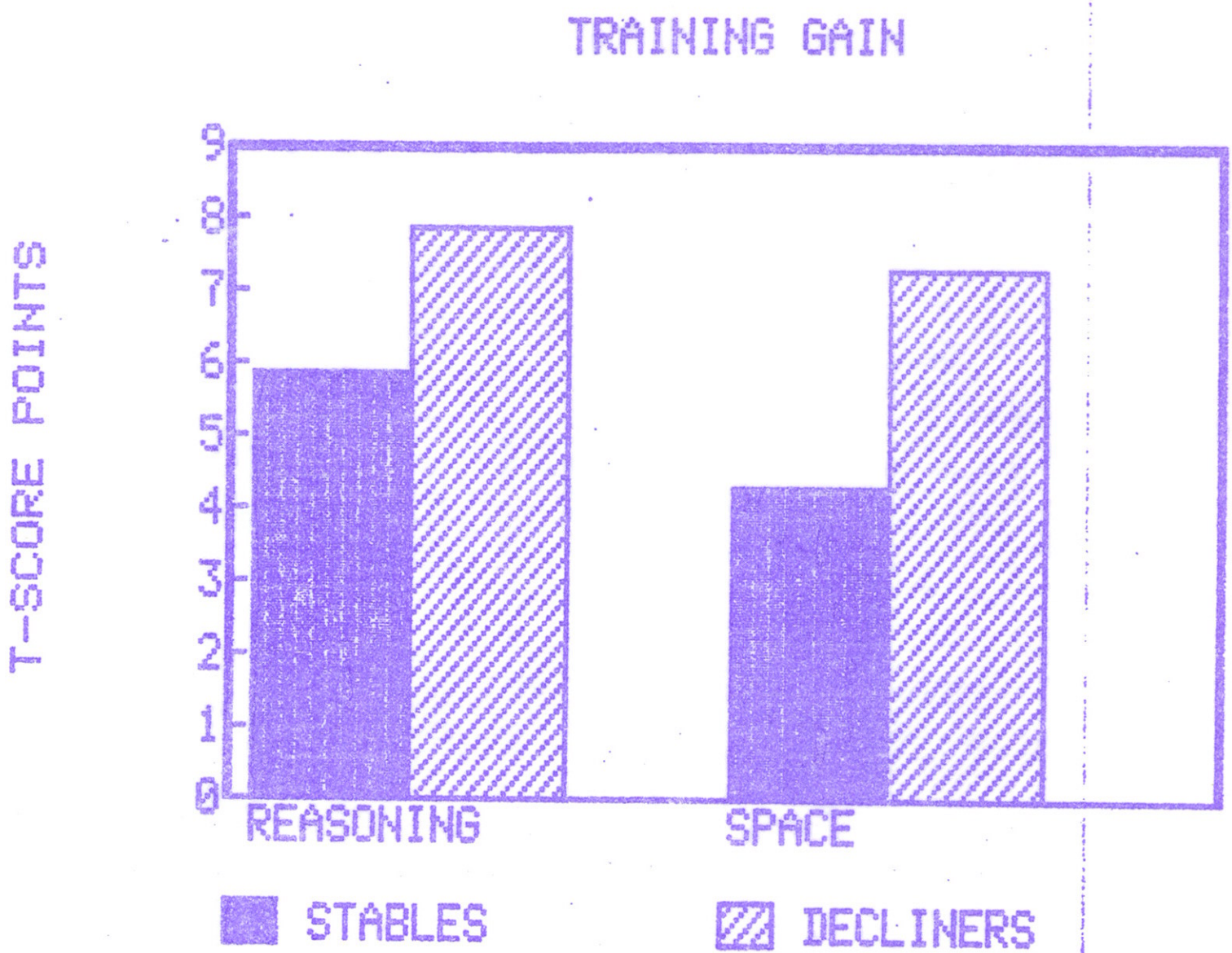
P M A SPATIAL ORIENTATION

5.

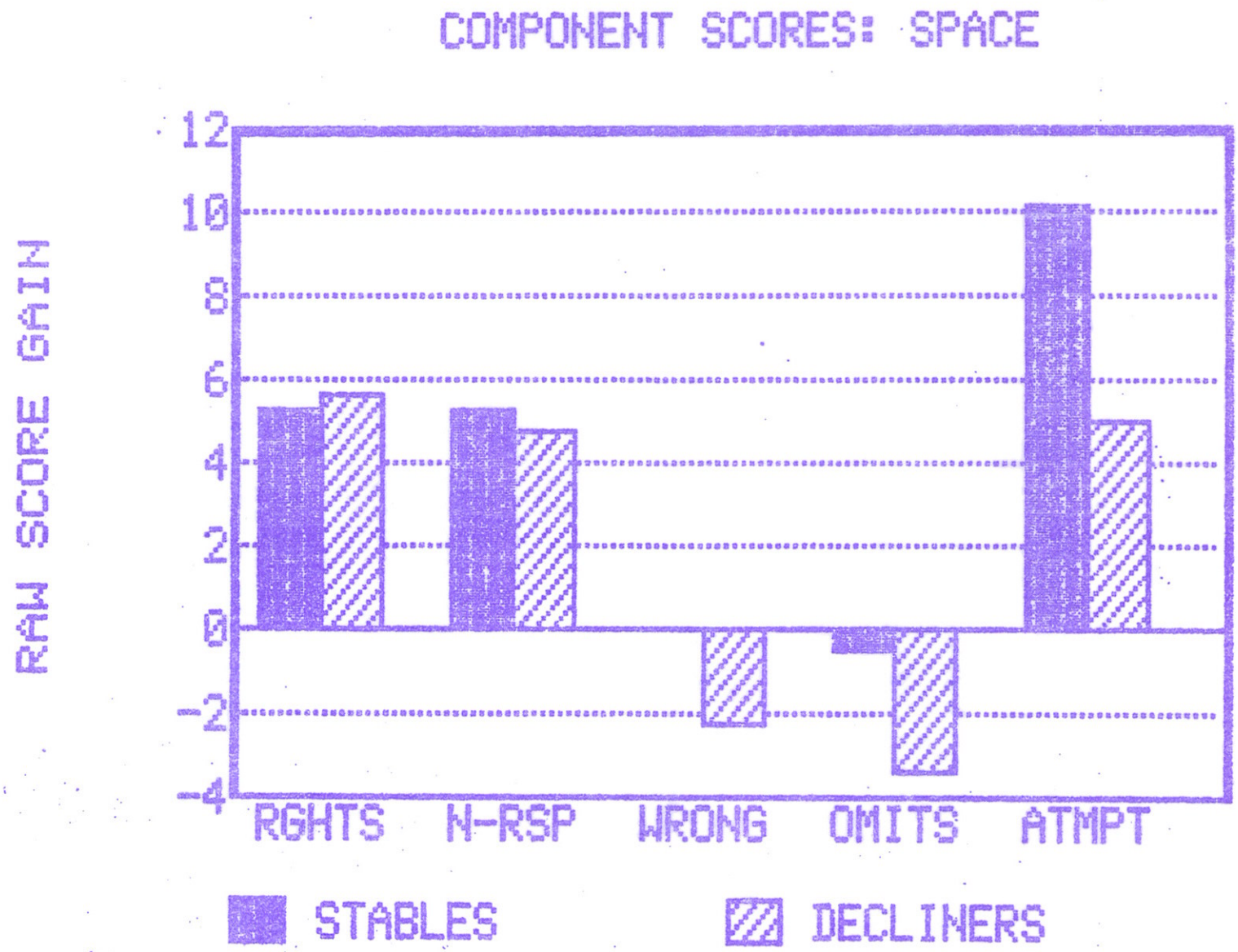


SPACE TRAINING-

- Task Analysis
- Concrete Labels for Angles
- Manual and Mental Rotation
- Use of Familiar Figures
- Subject-generated Labels for Abstract Figures
- Focusing on Two or More Features of Figure

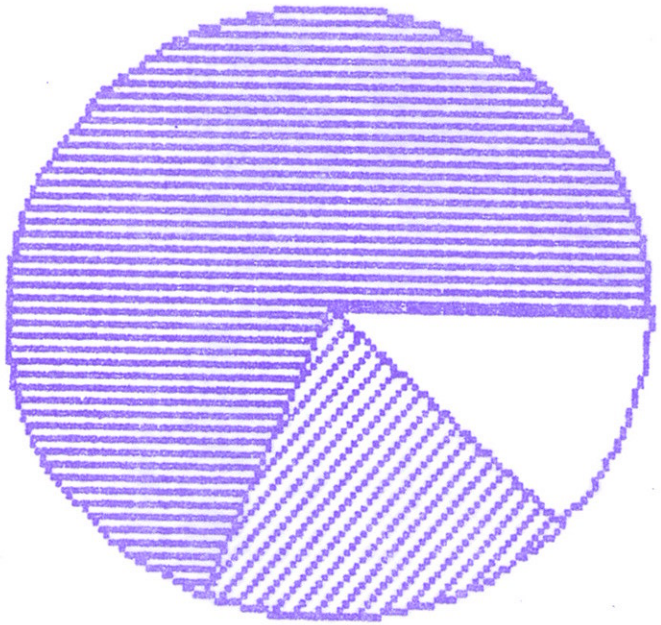

TRAINING GAIN
T-SCORE POINTS
9
8
7
6
5
4
3
2
1
0
REASONING
SPACE
STABLES
DECLINERS

% RETURNING TO 1970 LEVEL


COMPONENT SCORES: SPACE
RAW SCORE GAIN
12
10
8
6
4
2
0
-2
-4
RGHTS
N-RSP
WRONG
OMITS
ATMPT
STABLES
DECLINERS

GAIN IN CORRECT RESPONSES

PMA SPACE TRAINING



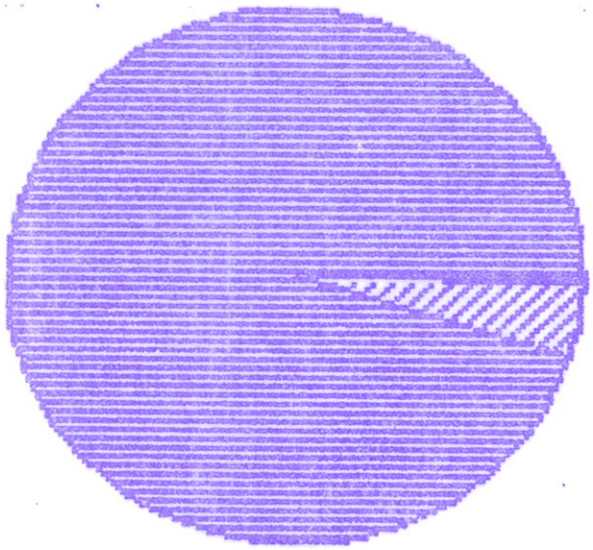
11.67%
LESS COMMISSION

19.47%
LESS OMISSIONS

68.86%
MORE ATTEMPTS

GAIN IN CORRECT RESPONSES

SPACE: STABLE SUBJECTS



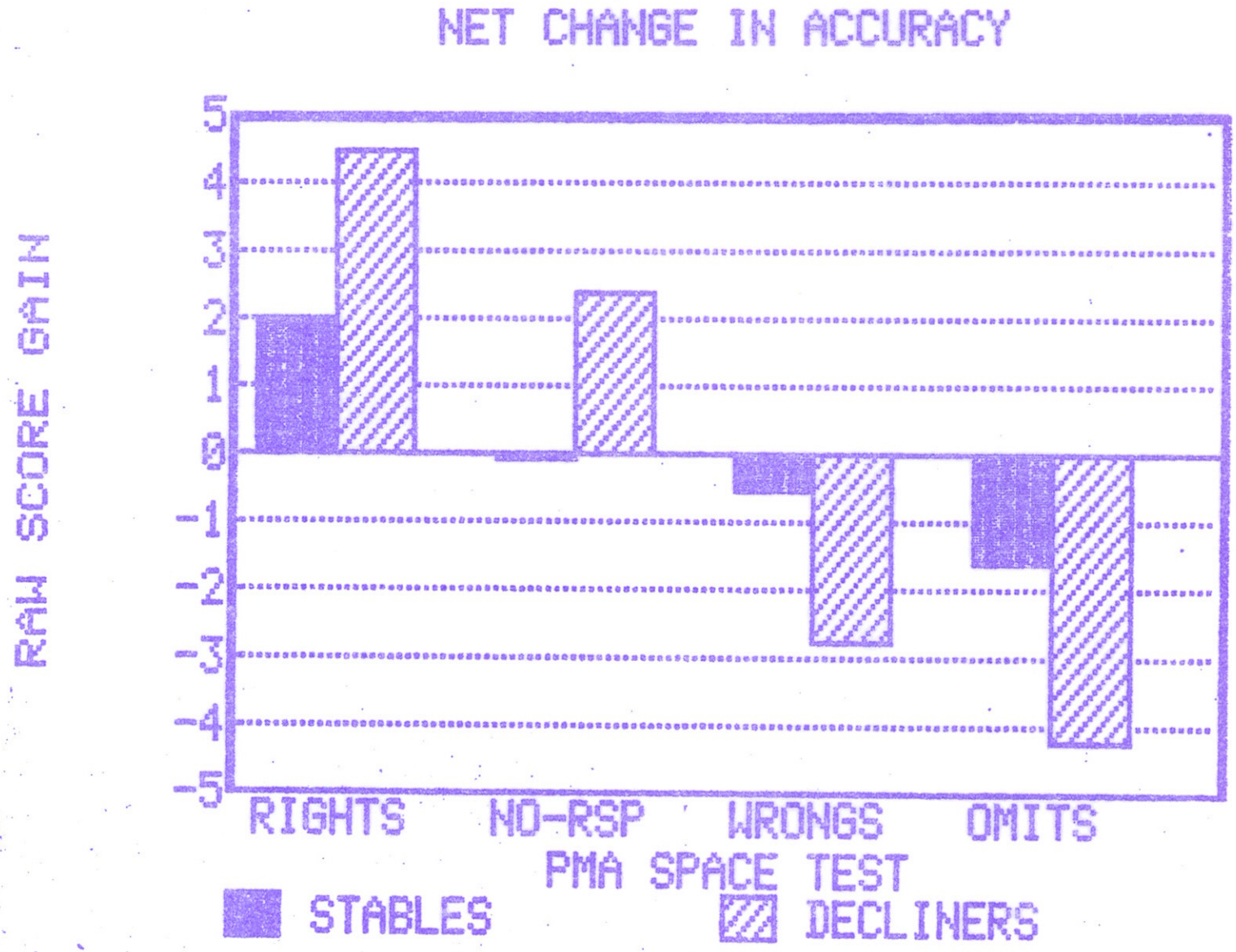
0%
FEWER OMITS

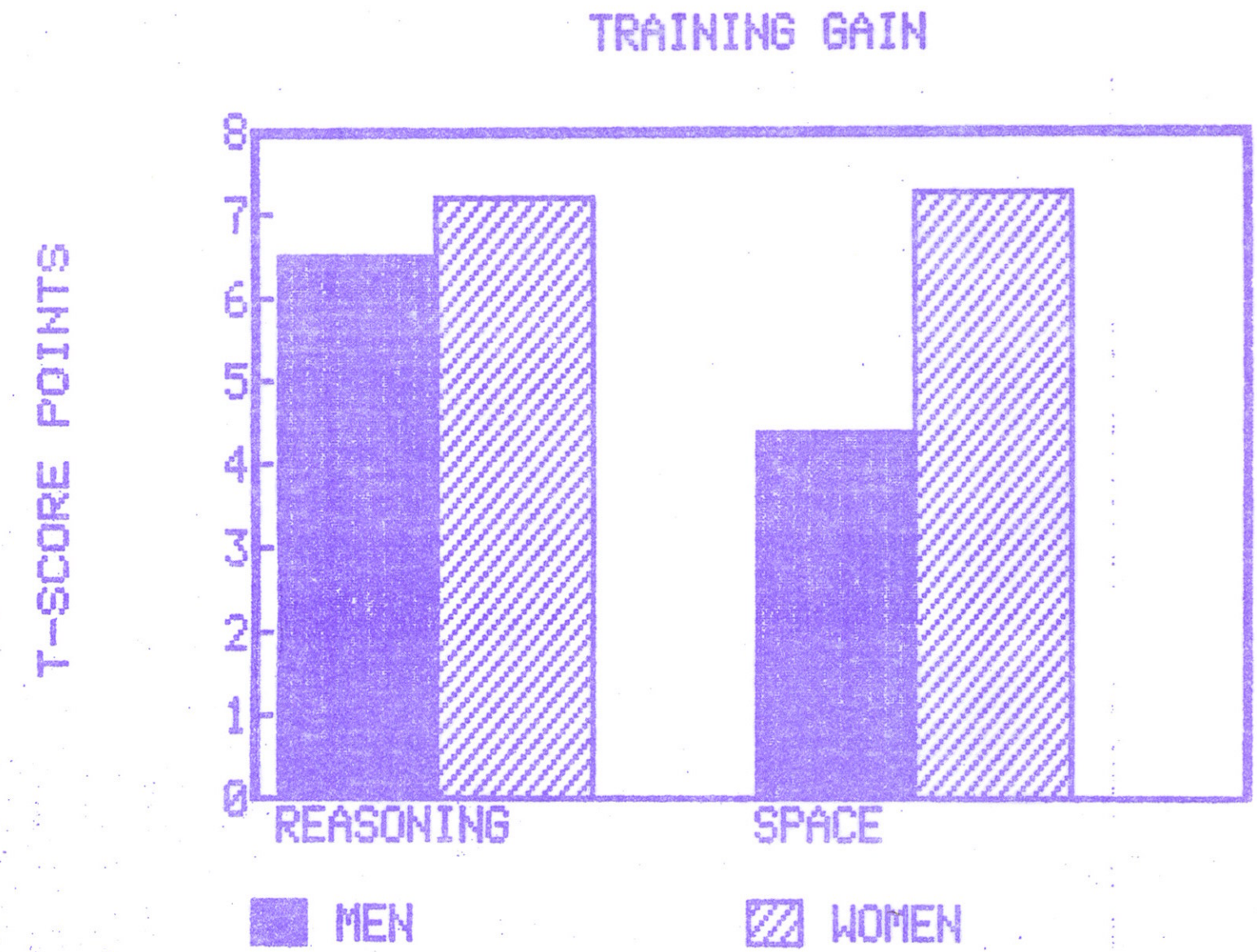
4.57%
FEWER COMITS

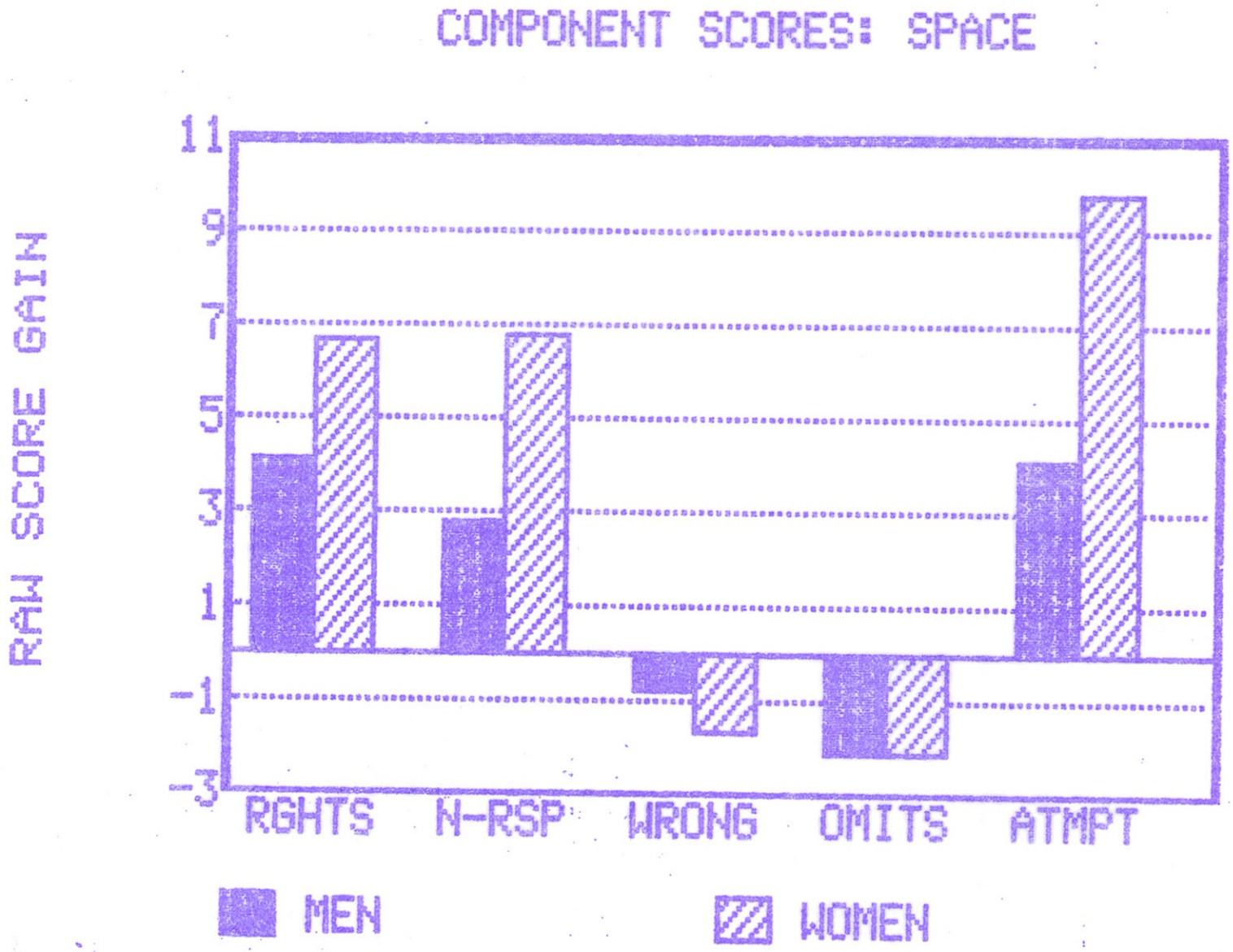
95.43%
MORE ATTEMPTS

SPACE: DECLINE SUBJECTS

20.65%
FEWER OMITS

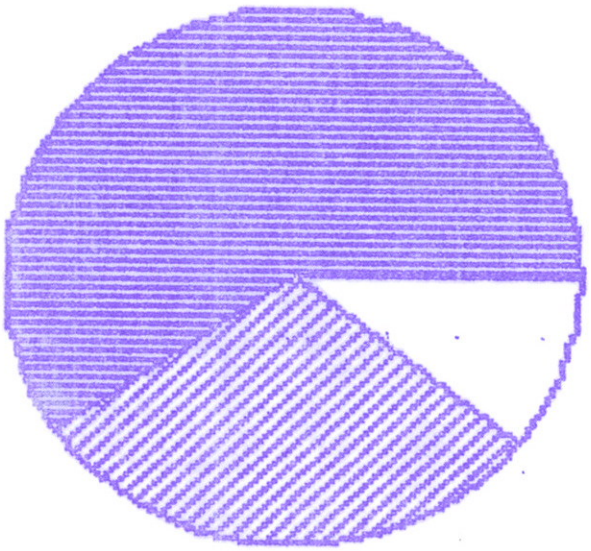

NET CHANGE IN ACCURACY
RAW SCORE GAIN
5
4
3
2
1
0
-1
-2
-3
-4
-5
RIGHTS
NO-RSP
WRONGS
OMITS
PMA SPACE TEST
STABLES
DECLINERS


TRAINING GAIN
T-SCORE POINTS
8
7
6
5
4
3
2
1
0
REASONING
SPACE
MEN
WOMEN


COMPONENT SCORES: SPACE
RAW SCORE GAIN
11
9
7
5
3
1
-1
-3
RGHTS
N-RSP
WRONG
OMITS
ATMPT
MEN
WOMEN

GAIN IN CORRECT RESPONSES

SPACE: MALE SUBJECTS

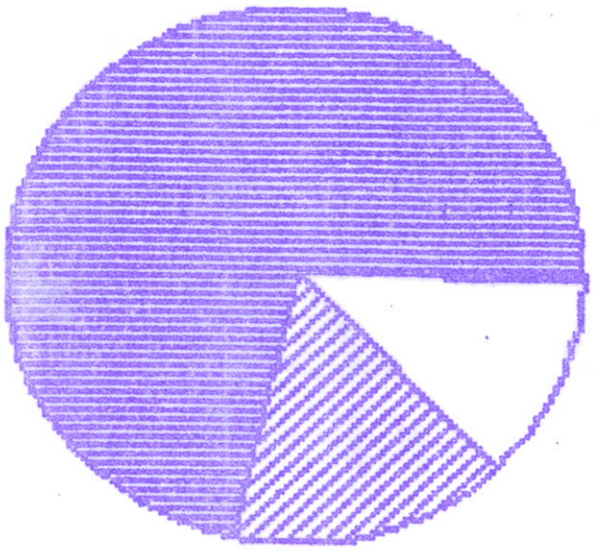


10.47%
FEWER COMITS

29.84%
FEWER OMITS

59.69%
MORE ATTEMPTS

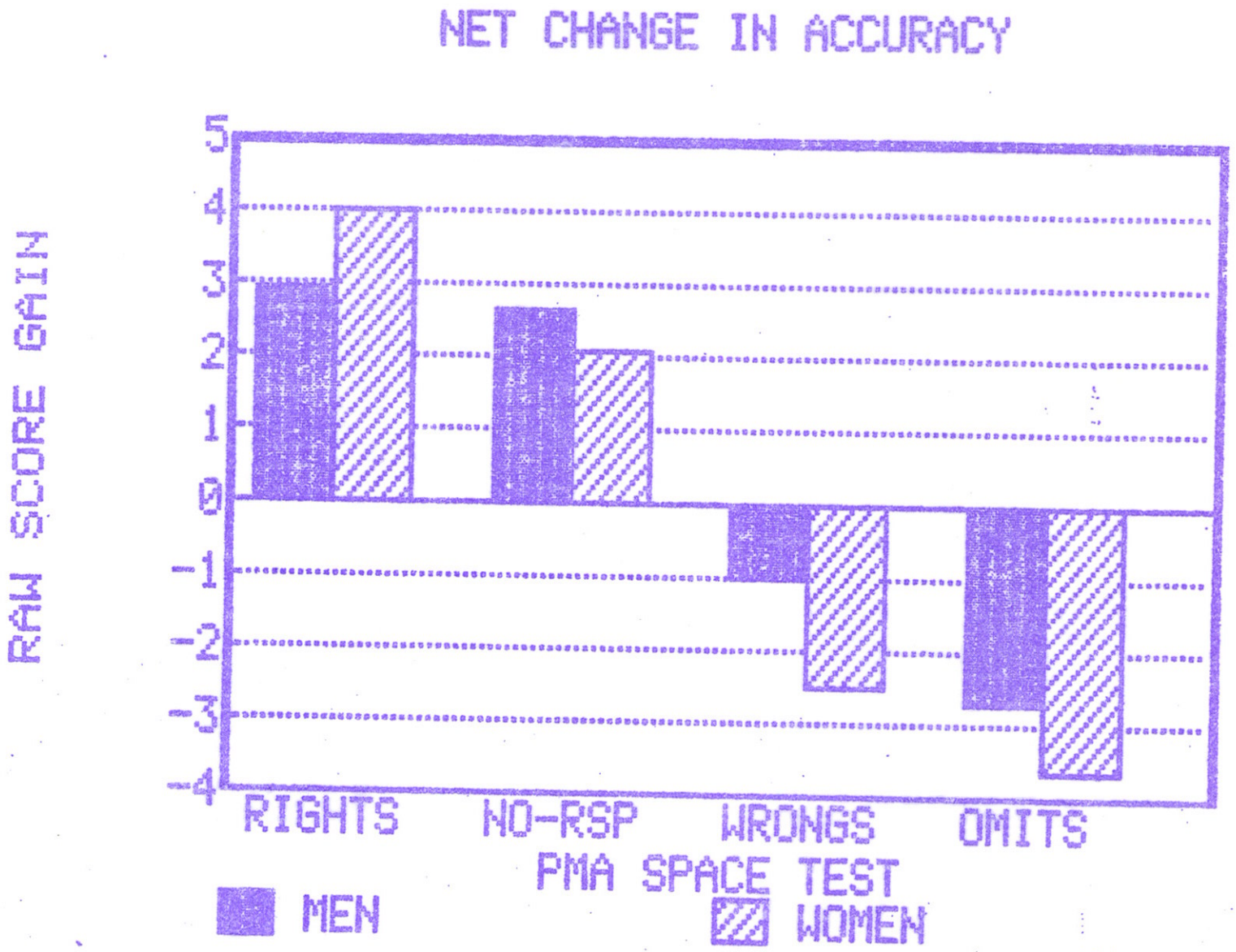
SPACE: FEMALE SUBJECTS



12.60%
FEWER COMITS

15.83%
FEWER OMITS

71.57%
MORE ATTEMPTS


NET CHANGE IN ACCURACY
RAW SCORE GAIN
5
4
3
2
1
0
-1
-2
-3
-4
RIGHTS
NO-RSP
WRONGS
OMITS
PMA SPACE TEST
MEN
WOMEN