1975

Training elementary school children to self-determine standards : effects on performance and generalization

Esther Anne Cohen
University of the Pacific

TRAINING ELEMENTARY SCHOOL CHILDREN TO SELF-DETERMINE STANDARDS: EFFECTS ON PERFORMANCE AND GENERALIZATION

A Thesis

Presented to

the Graduate Faculty of the

University of the Pacific

In Partial Fulfillment

of the Requirements for the Degree

Master of Arts

---

by

Esther Anne Cohen

August 1975

This thesis, written and submitted by

Esther Anne Cohen

is approved for recommendation to the Committee on Graduate Studies, University of the Pacific.

Department Chairman or Dean:

Thesis Committee:

Chairman

Dated July 11, 1975

ACKNOWLEDGEMENTS

This study would not have been possible without the assistance and cooperation of Mr. Ernest Tomlin, 6th grade teacher at Hoover Elementary School, and the 30 children in his class, who served as subjects for this study. In addition thanks are due to my thesis chairman, and committee members.

Abstract

The effects of instructions, non-contingent reinforcement, and contingent reinforcement on the level of self-determined standards for both experimental and generalization items were assessed using a reversal design. In addition the effects of the level of self-determined standards on performance in spelling was examined. The results indicated that contingent reinforcement for raising standards was an effective means of training the 30 subjects in this study to set high standards of performance for both experimental and generalization items. High standards did not, however, result in high levels of performance. When reinforcement was contingent upon subjects raising their standards and then scoring at this higher level (matching), both the level of standards and the level of performance showed moderate increases for most subjects. Subjects also exhibited similar increases in standards and performance for the generalization items. The maintenance of these increases during the subsequent reversal phase, however, limited the extent to which the changes could be attributed to the experimental manipulation. Further research is needed to evaluate more fully the effects of a matching procedure on the level of self-determined standards and performance.

Training Elementary School Children to Self-Determine Standards: Effects on Performance and Generalization

One of the primary objectives of the educational system is to teach children to be independent, to accept responsibility for their actions, and to respect the rights of others. Children are expected to learn how to behave in a socially acceptable manner in the absence of external controls; that is, they are expected to learn self-control.

Self-control can be conceived of as the factor necessary to maintain appropriate behaviors when external regulations and contingencies are withdrawn. It may be that self-control behaviors are necessary before changes in behavior can generalize to other situations or to other behaviors. In other words, programming the generalization of treatment effects may be accomplished by teaching self-control (O'Leary & Drabman, 1971). Unfortunately many educators have not been able to specify what behaviors constitute self-control, and as a result they have not been able to systematically teach these behaviors (Kanfer, 1973).

Behavioral research conducted in both applied and laboratory settings has been successful, in part, in identifying the behaviors that comprise self-control or self-management (Bandura, 1971; Glynn, Thomas, & Shee, 1973). These include self-monitoring, self-determination of standards, self-evaluation, and self-reinforcement. Although these behaviors have been studied individually, the nature of the relationship between them makes it difficult to isolate the

effects of each independently. For example, self-monitoring may result in self-evaluations based on a self-imposed standard. These self-evaluations may, in turn, generate self-reinforcing or self-punishing consequences (Bandura, 1971). Nevertheless, the specification of the components of self-management behavior has furthered the development of training procedures for these behaviors.

Research on self-management behaviors in classroom settings indicates that children are not by nature good managers of their own behavior. Broden, Hall, & Mitts (1971) conducted a study in which an eighth grade girl was asked to keep a record of the frequency of her study behavior during class. A reliability observer collected corroborative data on the behavior. The reliability between the child and the observer was low, suggesting that the child needed to be taught accurate self-monitoring. In another study, adolescent psychiatric residents were required to self-evaluate their behavior in the classroom (Santogrossi, O'Leary, Romanczyk, & Kaufman, 1973). They were to award themselves a rating from one to five, based on the amount of disruptive behavior they displayed. Students consistently gave themselves the highest rating regardless of their behavior, indicating that training in self-evaluation is necessary. In addition, both Lovitt & Curtis (1969), and Felixbrod & O'Leary (1973) conducted studies in which children were allowed to determine their own contingencies for reinforcement. That is, children decided how many problems they had to complete in order to receive one point. The results from both of these studies suggests that children tend to lower their performance requirements over time in order to maximize reinforcement.

This finding seems to indicate that children need to be taught how to set standards for themselves.

It appears that children require systematic training in self-management behaviors similar to the training they might receive for any other behavior. Furthermore the generalization of self-management behaviors must be programmed; that is, it cannot be expected to occur spontaneously (Baer, Wolf, & Risley, 1968). Thus far, procedures to teach accurate self-monitoring, self-evaluation, and self-administration of reinforcement have been developed (Bolstad & Johnson, 1972; Fixsen, Phillips, & Wolf, 1972; Drabman, Spitalnik, & O'Leary, 1973). The most frequently used training technique has been to reinforce students for matching their self-assessments with those obtained by the teacher. The maintenance of these various self-management behaviors has been produced by gradually fading the matching contingency and using social praise as a reward for appropriate classroom behaviors (Drabman, Spitalnik, O'Leary, 1973). What remains to be developed, however, is a procedure for teaching children how to set appropriate standards of performance for themselves in a classroom environment.

The research that has been conducted on the self-determination of standards thus far is somewhat confusing. This confusion can be attributed to the different ways in which the self-determination of standards has been conceptualized. On the theoretical level, the self-determination of standards has been viewed as both a component part of self-reinforcement behavior (Bandura, 1971), and as a behavior independent of self-reinforcement. On the experimental level, the

self-determination of standards has often been confused with the self-determination of reinforcement (Glynn, 1970; Lovitt & Curtis, 1969; Felixbrod & O'Leary, 1973). If an individual simply determines the magnitude of reward a particular behavior merits, before or after the behavior is completed, he is self-determining reinforcement but not standards. Standard setting must occur before the behavior is performed, evaluated, or provided with a consequence. Furthermore, evaluation and the subsequent consequences should be contingent upon attaining that standard. Thus, in the first case the level of reinforcement is manipulated and in the second case, the level of the standard is manipulated.

Several of the variables influencing the type of standard an individual chooses to self-impose have been investigated in laboratory settings. In the experiments conducted by Kanfer and his associates, the frequency of self-reinforcement, with respect to a task in which the accuracy and adequacy of performance was poorly defined, was assessed. This frequency measure of self-reinforcement was then used to determine whether subjects had imposed a lenient or stringent standard for performance. As a result of a series of experiments, the following variables have been identified as influencing the rate of self-reinforcement and hence, the level of self-determined standards: the type of instructions a subject receives regarding his criterion for self-reinforcement (Kanfer & Marston, 1963); the rate of externally administered reinforcement in a prior training condition (Kanfer & Duerfeldt, 1968); and the magnitude of the self-administered reward (Marston & Kanfer, 1963).

Bandura (1971) and his co-workers used a different strategy to investigate the variables affecting the acquisition of standards for performance. For their purposes, the self-determination of standards included instances in which an individual adopts standards set by some external agent, as his own, instances in which he adopts a standard for himself relative to some modal level of performance, and instances in which he determines his standard by means of a social comparison process (Bandura, 1971). Bandura incorporated these aspects of standard setting in his investigations by studying the role of several modeling processes on the type of standard a subject self-imposed. In these studies subjects usually observed one or more models performing a motor task in which they self-determined performance standards and self-administered rewards. Subjects were then required to perform the same motor task. The effects of the modeling conditions were determined by the degree to which subjects self-imposed the standards they had observed. It appears that consistency in the behavior of the models observed, and similarity between the model and subject with respect to competency at the task are two factors that increase the likelihood that a subject will adopt the standards of the model he observed (Bandura & Whalen, 1966; Bandura, Grusec, & Menlove, 1967).

In an applied study conducted in a classroom setting, Felixbrod & O'Leary (1973) identified several other variables that influence the type of performance standard an individual self-imposes. They specifically investigated the effects of the passage of time on the level of self-determined standards. Subjects self-selected a standard

of performance by determining the number of problems they would have to complete in order to earn one point, with respect to a 10 problem assignment. That is, the lowest standard that a subject could self-impose would be to make one problem worth one point, whereas the most stringent standard would be to make 10 problems worth one point. Subjects were allowed to change their standards for 5 out of a total of 6 sessions. The results indicated that subjects lowered their performance requirements over time. The authors also postulated that the degree of external control perceived by the subjects, and the absence of aversive consequences for lowering standards, both indirectly related to time, may affect the level of self-determined standards.

In summary, research in the area of the self-determination of standards has served to identify many of the variables that influence the type of standard an individual will self-impose. What remains to be investigated, however, is the possibility of using these variables to develop a procedure for teaching children to self-determine high standards for performance. Furthermore, the efficacy of such a training procedure must be viewed in terms of a child's ability to generalize from the training situation. In other words, if a child is trained to adopt high standards for performance in one area or on some items, he should also be able to adopt high standards in another area or on some similar items without additional training. As Premack & Anglin (1973) note, "a child demonstrates self-management when his behavior conforms to certain rules not only in the training situation but in the world at large." Thus, self-management cannot be said to have occurred unless

generalization of standard setting to other similar behaviors or in other similar situations occurs.

Another aspect of the self-determination of standards that deserves consideration is the effect of the level of self-determined standards on performance. Bandura & Perloff (1967) conducted a study in which children set their own performance contingencies for self-reinforcement on a motor task. Their findings suggest that self-imposed standards will result in high rates of behavior and that some children raise standards when given the opportunity. On the basis of these results and related research that has been conducted in the area of goal setting and achievement (Locke, Cartledge, & Koeppel, 1968), Bandura proposed that there is a direct relationship between the standard that is self-imposed and achievement; that is, high standards result in high attainments. He also suggested that once an individual had attained a certain level of performance he would no longer be satisfied with it, and would subsequently raise his standards. The proposed relationship between self-determined standards and achievement remains to be demonstrated in applied settings over time.

Studies that have been concerned with the effects of self-determined standards on performance in classroom settings suggest only that self-imposed standards are as effective as externally imposed contingencies with respect to maintaining effortful behavior for short periods of time (Glynn, 1970; Lovitt & Curtis, 1969). These findings are open to question, however, because the individual effects of self-determined standards were confounded with the effects of other

self-management behaviors and external regulations which were also controlling performance.

Based on the research that has been conducted in the area of self-management in the classroom thus far, it can be concluded that children need to be taught how to manage their own behavior; that training procedures have been successfully developed for several of the components of self-management behavior; that a training procedure for teaching children how to set their own standards for performance is needed; that the effects of such a procedure should generalize to other similar situations in order to be truly effective; and finally that the effects of the level of self-determined standards on performance warrants further study. The present study was designed to (1) assess the effects of instructions, non-contingent reinforcement, and contingent reinforcement for raising standards for performance on the level of self-determined standards, (2) to determine whether subjects who have been trained to self-impose stringent standards of performance for one set of items on an academic task will also self-impose those standards on another set of items without additional training, and (3) to determine the effects of the level of self-determined standards on academic achievement.

Method

Subjects and Setting

Thirty-five sixth-grade pupils from an elementary school in West Stockton served as subjects in this study. One subject was eliminated from the study because she achieved the maximum level of

performance throughout all phases of the study. Four additional subjects were eliminated due to excessive absences. Thus, data are reported for 30 out of the 35 children.

The experimenter and the teacher conducted the study in the classroom for approximately 20 minutes each day. Upon the experimenter's arrival, the teacher would stop the on-going class activity to allow the students to participate in the study. After the experimental session was over, the teacher would resume the regular classroom activity. The study did not in any other way interfere with the regular sixth-grade curriculum. The experimental sessions did not take place at the same time each day.

Materials

Subjects received a spelling list of 30 words every other day (see Appendix A). In order to control for the level of difficulty between spelling lists, 30 words were randomly selected each time, from a combined sixth- and seventh-grade spelling list of 2,485 words. No word could be chosen more than once. Fifteen words were then randomly chosen from the 30, and served as the words for which the students could earn privileges (privilege words). Stars (asterisks) were always placed by these words to distinguish them from the words for which no consequences were provided (probe words). When the subjects were tested on the spelling words, they received a test paper numbered from one to thirty with a star placed next to the number of a word which was a privilege word. In addition, at the top of every spelling list and spelling test paper was a symbol code explaining that a star signified

a privilege word. In this way subjects could readily discriminate between the words for which they could earn privileges and the words for which they could not.

Subjects also received a standard sheet upon which they recorded their standards; that is, they recorded the number of spelling words they felt they should spell correctly in order to receive a performance rating of excellent, good, satisfactory, and unsatisfactory, for both the privilege words and the probe words (see Appendix B). The standard sheet was divided into four columns, one for each of the performance ratings, and two rows. One row, labeled "privilege words," was for the standards the subjects set for the privilege words. The second row, labeled "other words" was for the standards the subjects set for the probe words. The standard sheet format was altered slightly depending upon which of the experimental conditions was being implemented.

Response Definitions and Measurement

The number of correctly spelled privilege words that a subject considered appropriate for a rating of excellent work, good work, satisfactory work, and unsatisfactory work served as the measure of the level of self-determined standards. The number of correctly spelled probe words that a subject considered appropriate for a performance rating of excellent, good, satisfactory, and unsatisfactory provided a measure of generalization with respect to standard setting.

The level of performance was measured by the number of privilege words spelled correctly on a spelling test. The number of probe words spelled correctly on a spelling test constituted the measure of

generalization with respect to the level of performance.

Measures of the level of self-determined standards were taken two or three times weekly. Reliability checks on the experimenter's recording of these data from the standard sheets to data summary sheets were taken once during each phase, and equalled 100%. Measures of the level of performance were taken twice weekly. The students conducted reliability checks on the experimenter's grading of papers for each test that was given; that is, when they received their graded test paper they were required to check their work to insure that the papers had been scored correctly. Reliability on this measure also equalled 100%.

Procedures

Pre-training. Before baseline measures were taken, subjects participated in two training sessions. During one of these sessions subjects were given a list of classroom privileges to rank in order of preference. The ratings that the subjects assigned to a particular privilege were averaged and the privileges were then scaled on the basis of their average rank. In this way the privileges were ordered according to their desirability, as determined by the entire class. Privileges with the highest ranks were used as backup reinforcers for the point system that was implemented during the reinforcement phases and as reinforcers for the matching phase (see Appendix C).

The remaining privileges were assigned to each of the ratings of performance on the basis of their rank. Thus, the most desirable privileges were available only to students who had obtained a rating of excellent; privileges with the next highest rank were available to

students who had obtained a rating of excellent or good; privileges with the lowest ranks were available to students who had obtained a rating of excellent, good, or satisfactory (see Appendix D). Privileges were assigned to the performance ratings in this fashion to increase the probability that students would attach appropriate values to each of the ratings. For example, given the number and type of privileges a student could obtain with a rating of good, a rating of excellent should be more desirable to the student. Furthermore, the assignment of privileges to each of the ratings more closely simulates conditions under which students tend to maximize reward at the expense of performance (Kanfer, 1973). However, it was made clear to the children that there were no privileges available for obtaining any of the performance ratings for the 15 probe words.

During the second training session subjects were able to familiarize themselves with the procedure for setting their own standards. The experimenter presented several examples of how one might set his/her standards for 15 spelling words. The subjects then wrote down the number of words they felt they should spell correctly in order to obtain a rating of excellent, good, satisfactory, and unsatisfactory. They were not given any spelling words at this time.

Baseline. Subjects received their standard sheets and a list of privileges which could be obtained by acquiring a certain rating of performance on the privilege spelling words. The following instructions were then given:

> Soon I will hand you a list of 30 spelling words; 15 will have stars and 15 will not. Tomorrow you will be tested

> on these words. You will receive either an excellent, good, satisfactory, or unsatisfactory for a grade depending upon the standards you set for yourself. You will receive two grades, one for the words with stars and one for the other words. Now look at the list of privileges. You may only earn privileges for your grade for the words with stars. So, if you got an excellent as a grade for the words with stars or privilege words you could choose any privilege from the list. It doesn't matter what grade you get for the other words; you may only choose privileges based on your grade for the privilege words.
>
> Now look at your standard sheets and decide how many words you think you should spell correctly in order to get an excellent, a good, a satisfactory, and an unsatisfactory for the privilege words. Then in the space marked other words do the same thing for the other 15 spelling words.

Thus, subjects had to set their standards twice, once for the privilege words and once for the probe words. The standard sheets were collected and the spelling word lists were distributed. The following day the experimenter tested the subjects on the 30 words in the usual fashion; that is, she stated the word, used it in a sentence, and then repeated it. The subjects used a special spelling test paper on which to write down the words, so that they could distinguish between the privilege words and probe words. (See materials.)

The experimenter graded the papers, and circled the number of words spelled correctly at the top of the page, followed by the appropriate performance rating as determined by the subject's standards. This was done for both the privilege and probe words. The grades were also recorded under the appropriate rating on the standard sheet. The tests were returned the next day, at which time the subjects were instructed to check the grading of their papers. Any necessary corrections were made. A privilege sheet was circulated around the classroom upon which subjects could sign up for the privilege they

wanted, according to the performance rating they had received for the privilege words. A volunteer privilege monitor kept a record of which subjects had used a particular privilege and was responsible for reminding the teacher to present opportunities for the privileges to be used.

Subjects would then set their standards again, and receive a new spelling list. Thus, subjects were able to set their standards as high or low as they wanted, and could change their standards as often as they wanted. Their standard sheets provided them with a complete record of the level of their standards and performance for all the preceding spelling tests. Baseline lasted for 10 days.

Instructions. The procedures during this phase were essentially the same as baseline, except that before the subjects set their standards, the experimenter repeated the following instruction: "Good students should set high standards for the privilege words." This phase was instituted for 4 days to examine what effect instructions alone might have on the level of self-determined standards.

Non-Contingent Reinforcement. Subjects received five points each time they set their standards during this condition. The points were recorded on a new standard sheet next to the row labeled "privilege words" (see Appendix B). The points could be exchanged for privileges that were listed on a special privilege sheet (see Appendix C). The following instructions were given each day before the subjects set their standards:

> In order to help you try harder on your privilege words I am giving everyone five points every time you set your standards. You can save your points or spend them right away. The special privilege list tells you how many

points each privilege costs.

The purpose of this condition was to evaluate the effects of "incentives" that parents and teachers often use to try to obtain desirable performance from their children. The non-contingent reinforcement phase lasted for 4 days.

Contingent Reinforcement for Raising Standards. Subjects were told that they would no longer receive five points every time they set their standards. They were allowed to exchange their remaining points and then received new instructions explaining how they might now earn points in order to obtain the special privileges.

> In order to earn points you must raise your standards for the privilege words. If you raise your standards for one of the ratings (excellent, good, or satisfactory) you will earn one point; if you raise your standards for two of the ratings you will earn two points; and if you raise your standards for all three you will earn five points. If you select the highest standards possible (15 for excellent, 14 for good, 13 for satisfactory) as your standards you will earn 10 points. You will lose the same amount of points if you lower your standards for one, two or all three of the performance ratings. These points can only be earned or lost if you raise or lower standards for the privilege words; not if you raise or lower your standards for the other words.

The response cost procedure (losing points for lowering standards) was included in this eight day reinforcement phase in order to discourage subjects from lowering and then raising their standards as a means of earning points. Points were recorded and exchanged in the same fashion as in the non-contingent reinforcement condition. Written instructions explaining how to earn points were also given to each subject (see Appendix E).

Reversal 1. In order to confirm the effects of contingent reinforcement on the level of self-determined standards, a return to baseline condition was implemented for 6 days. Subjects received a standard sheet of the same format as during baseline, and were told that they would no longer receive points for raising standards nor lose points for lowering standards. They were allowed to exchange their remaining points and received the same instruction as in baseline which was:

> I want you to decide, as you have been doing, how many words you think you should spell correctly in order to get a rating of excellent, good, satisfactory, and unsatisfactory for both the privilege words and the other words.

Matching standards and performance. During this phase subjects had to select the rating of performance they wanted to receive on a particular spelling test, and then had to obtain that rating in order to be able to earn privileges for their performance on the privilege words. Subjects received a new standard sheet (see Appendix B), which provided a space for them to record the performance rating they wished to obtain on each spelling test.

In addition the privilege sheet was revised so that the more desirable privileges were now available for obtaining the different ratings of performance (see Appendix C). That is, the privileges that had formerly been reserved for the point system in the reinforcement phases were now incorporated into the regular list of privileges. The desirability of each privilege was determined by the average rank the privilege has received during the pre-training session and by the number of children who had picked that privilege in the previous phases. The privileges were then assigned to each of the performance ratings

accordingly. The privilege sheet was revised in an attempt to make obtaining a particular rating of performance (meeting one's goal) even more reinforcing, and to make obtaining a higher rating of performance (raising one's goal) even more reinforcing, than in the previous phases. Subjects were also able to earn certain "bonus" privileges by attaining their preselected rating three times in succession. These bonus privileges were included in order to promote consistency in the subjects' level of performance.

The subjects received written instructions explaining the new procedures (see Appendix F) in addition to the following verbal instructions:

> In order to be able to pick a privilege from the privilege sheet you must first pick a goal. That is, you must pick a rating that you want to get on the next spelling test (excellent, good, or satisfactory), and then you must get that rating in order to pick a privilege. So if you decide you want to get a good on the privilege words, you must get a good in order to pick from the privilege sheet. You may pick the same performance rating as your goal, only twice. Then you must either choose a higher rating as your goal or raise your standards for excellent. Once you have reached your goal twice you may not pick a privilege again until you pick a higher goal and then meet that goal. You will also be able to pick a bonus privilege if you meet your goal three times in a row. This is only for the standards you set for the privilege words.

Thus in this condition subjects could only receive privileges if they raised their standards and then were able to meet the increased performance requirements, or if they maintained standards at the maximum level and were able to perform at that level. This matching procedure was instituted for 10 days as a means of teaching the children to raise standards and consequently improve their level of performance.

Reversal II. This second return to baseline condition was conducted for 3 days in order to assess the effects of the matching condition on the level of self-determined standards and the level of performance. Subjects were told to continue setting their standards but that they no longer needed to pick a goal for the next spelling test. Standard sheets and the list of privileges were the same as in the original baseline phase. Subjects could once again earn privileges solely based on the performance rating they obtained.

Results

Group Analysis

Figure 1a represents the level of self-determined standards for all subjects for the three ratings of performance (excellent, good, satisfactory) across all conditions. Figure 1b represents the mean standards for all subjects across all conditions, and Figure 1c represents the number of words spelled correctly (level of performance) for all subjects across all conditions. The mean standard is an average of the three standards set for the individual ratings of performance. This figure will be used to describe the changes in the level of self-determined standards that occurred. Data were combined across days such that each data point represents two experimental sessions. Data are presented for both experimental and probe words. In addition, Table 1 presents the number of subjects who increased, decreased, or maintained their level of self-determined standards and their level of performance during the treatment and reversal phases.

Table 1

Number of Subjects Who Changed Standards and Performance During the Treatment and Reversal Phases

Direction of Change	Conditions			
	Contingent Reinforcement	Reversal	Matching	Reversal
Standards				
Increased	30	0	15	0
Decreased	0	17	0	7
No Change	0	13	12	19
Other	0	0	3	4
Performance				
Increased	20	7	22	9
Decreased	2	11	2	8
No Change	8	12	6	13
Other	0	0	0	0

Note. Subjects who lowered standards or performance during either of the reversal phases, but remained at a level that exceeded baseline or reversal I levels, are accounted for in the no change column.

Figure 1a. Standards for each of the 3 ratings of performance

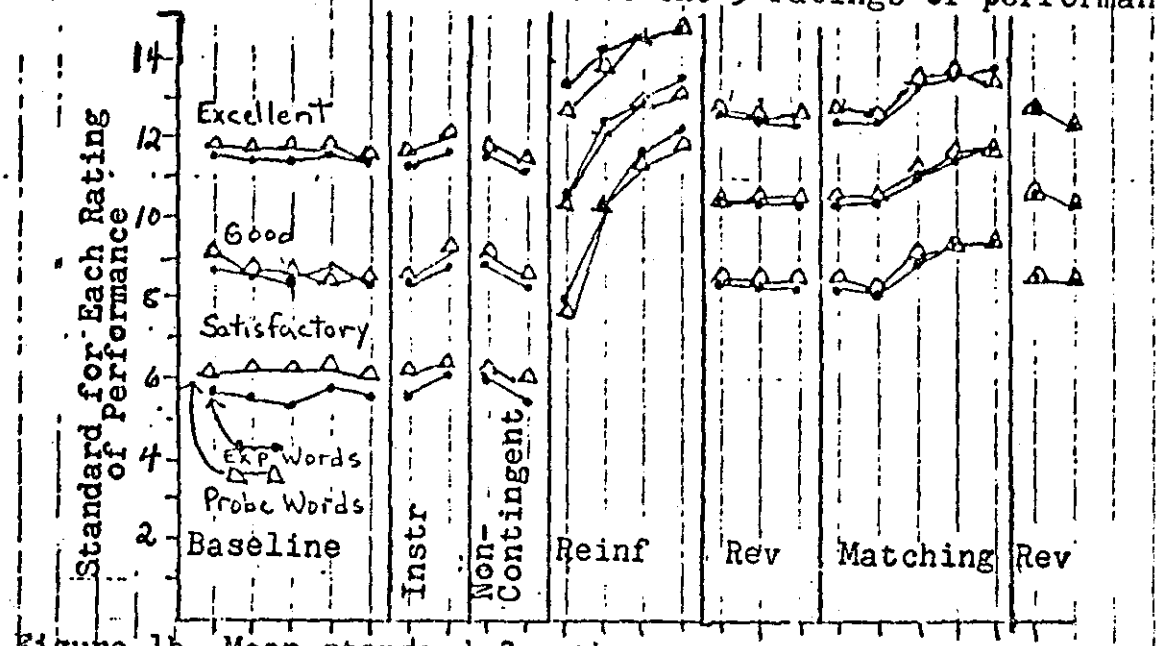


Figure 1b. Mean standard for the group.

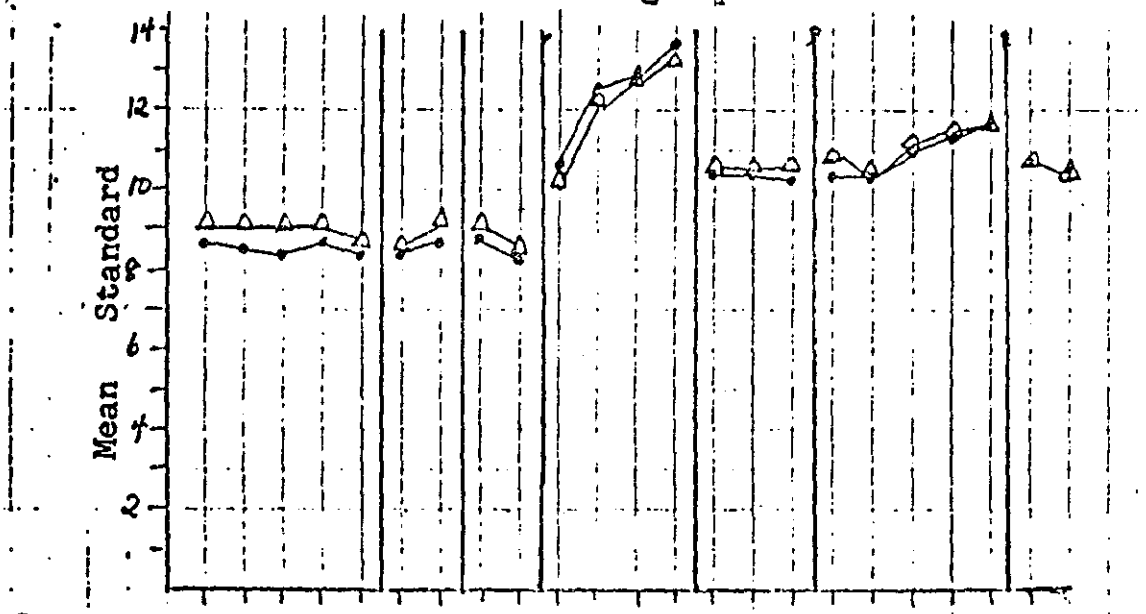
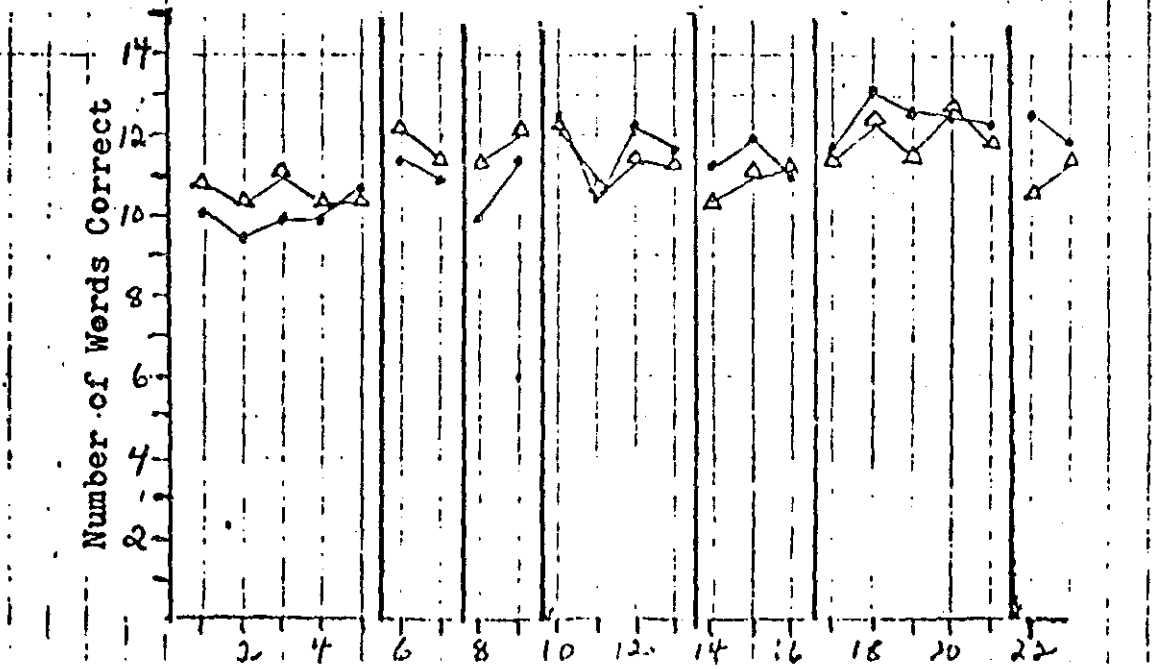


Figure 1c. Level of performance for the group



The group mean standard for the first three experimental conditions (baseline, instructions, and non-contingent reinforcement) remained relatively unchanged and equalled approximately 8.63 for the experimental words and 8.95 for the probe words. The group averaged 10.3 experimental words and 11.17 probe words correct during these three phases. Performance appeared to be increasing slightly across days (+.5).

During the contingent reinforcement phase the mean standard for both the experimental and probe words increased steadily and approached maximum (14) by the end of the phase. The mean number of words correct increased slightly and equalled 11.66 for the experimental words and 11.49 for the probe words.

The mean standard decreased during the first reversal phase to approximately 10.4 for the experimental words and to 10.73 for the probe words. The mean standard for both sets of words remained about 1.5 points above baseline. The average number of experimental words spelled correctly was 11.4 and the average number of probe words spelled correctly was 10.89, for the group.

During the matching condition, in which reinforcement was contingent upon increases in both standards and performance, the mean standard increased by about 2 points to almost 12 for both the experimental and probe words. The group averaged approximately 12.44 experimental words correct during this phase representing an increase of 1 point from the mean performance level in the reversal I phase. The average number of probe words correct equalled 12, also representing an increase of 1 point from the average obtained in the reversal I phase.

In the second reversal phase, the mean standard showed a decline to approximately the same level as reversal I for both the experimental and probe words (10.6 and 10.75 respectively). The average number of experimental words spelled correctly remained essentially unchanged (12.24), but the average number of probe words spelled correctly decreased slightly to 11.1.

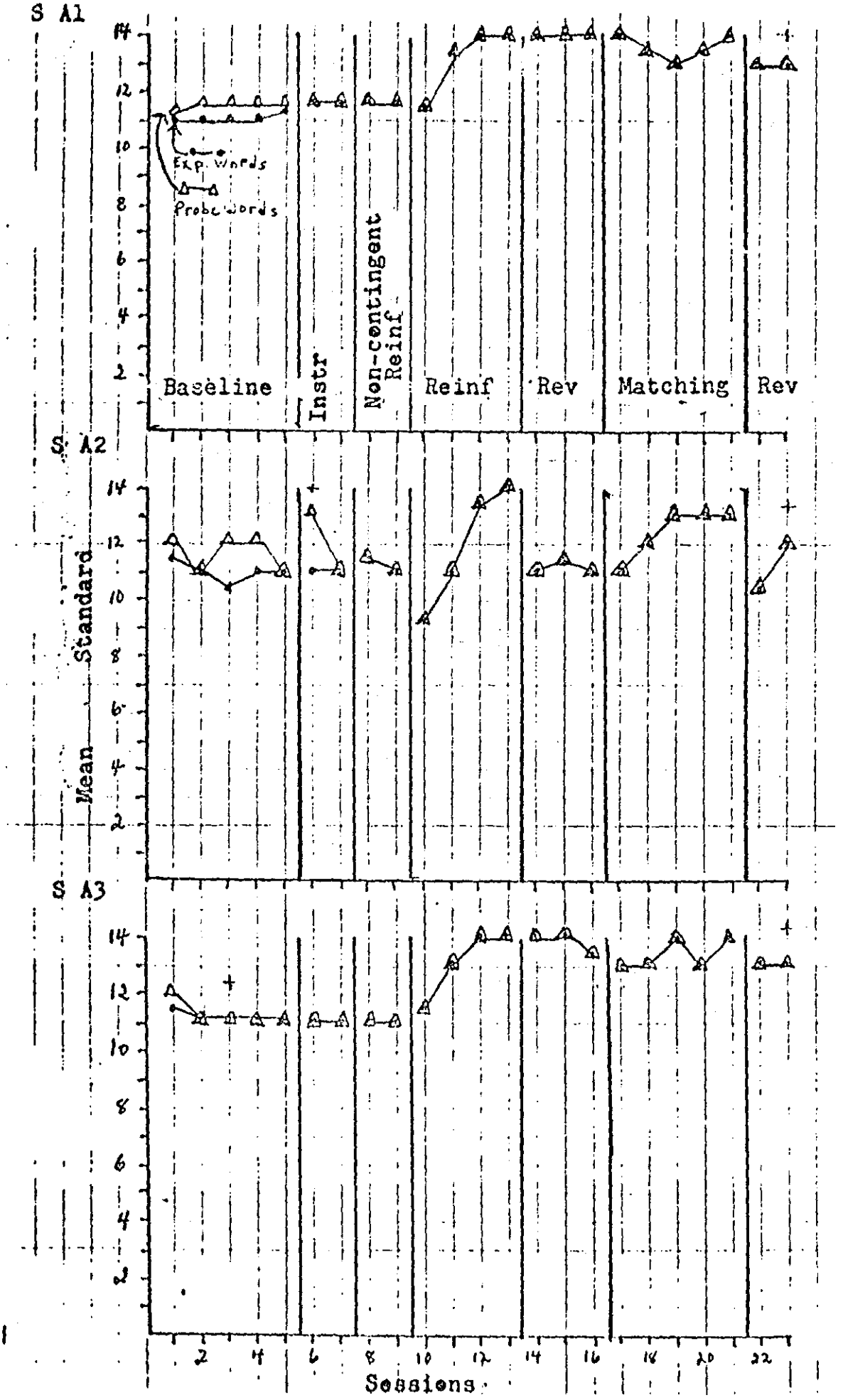
The mean standard and the level of performance for the probe words appear to parallel the mean standard and the level of performance for the experimental words in each phase. There exists, however, a slight discrepancy between the mean performance level for the probe words and the mean performance level for the experimental words during the reinforcement phases.

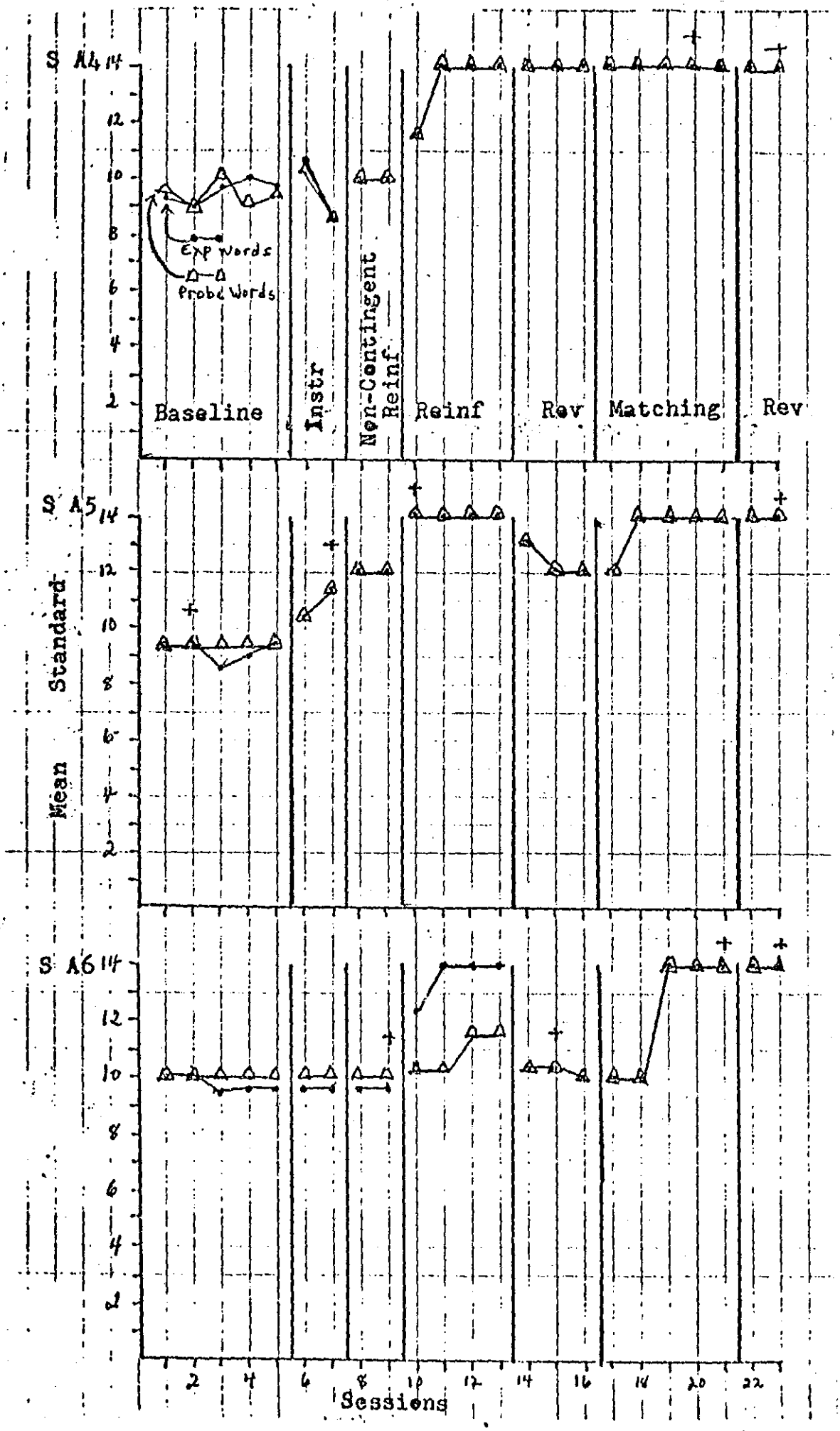
Single-Subject Analysis

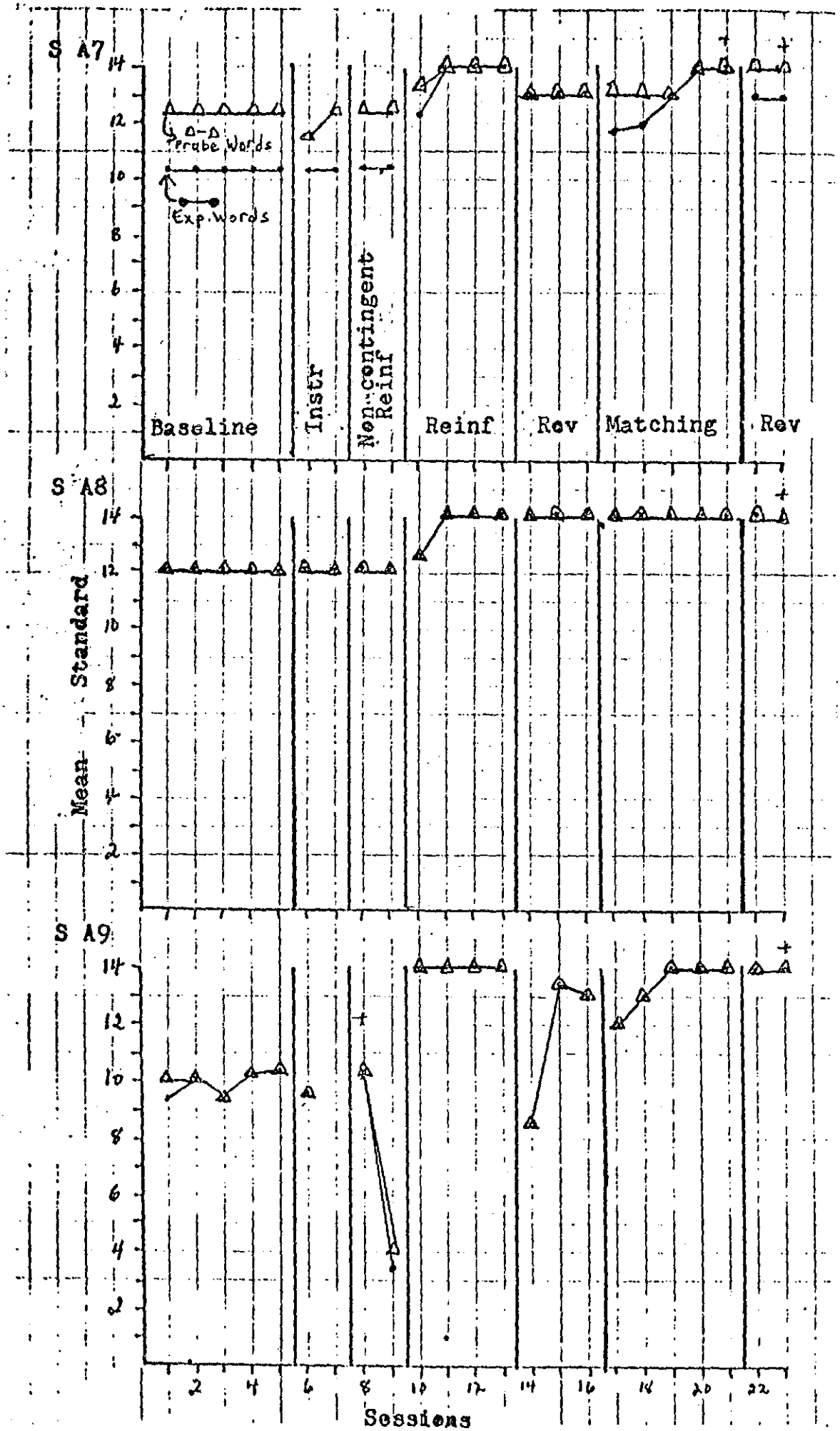
A single-subject analysis of the data was also performed in order to assess the degree to which individual subjects behaved in a manner suggested by the group data.

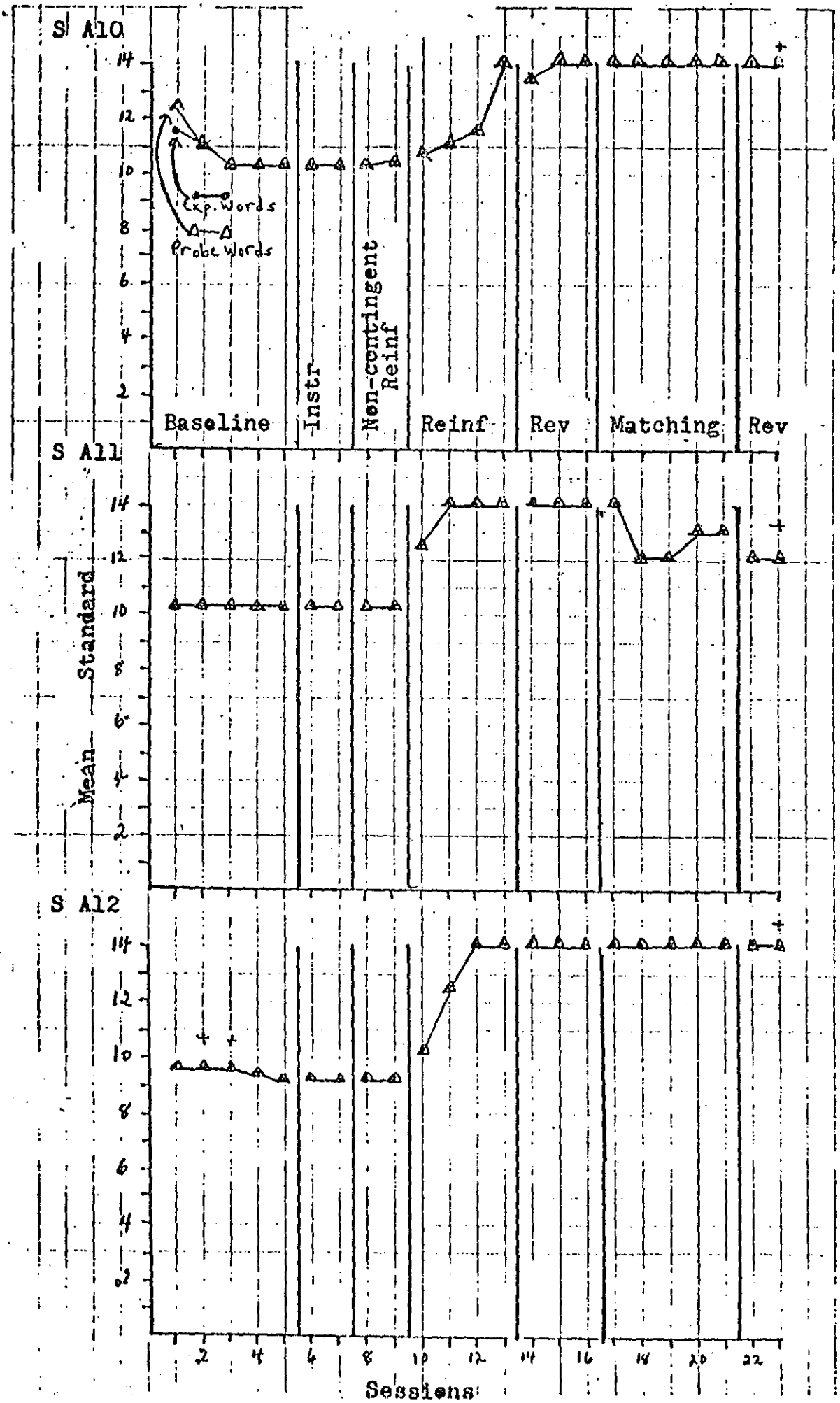
Standards. Figure 2 represents the mean standard for the experimental and probe words set by each subject across all conditions. The mean standard is an average of the standards set by a subject for the three ratings of performance (excellent, good, and satisfactory). The mean standard is approximately equal to the standard set by each subject for the performance rating of good. Data were combined across days such that each data point represents two experimental sessions. A plus mark (+) above a particular data point signifies the days upon which a data point represents a single experimental session. The

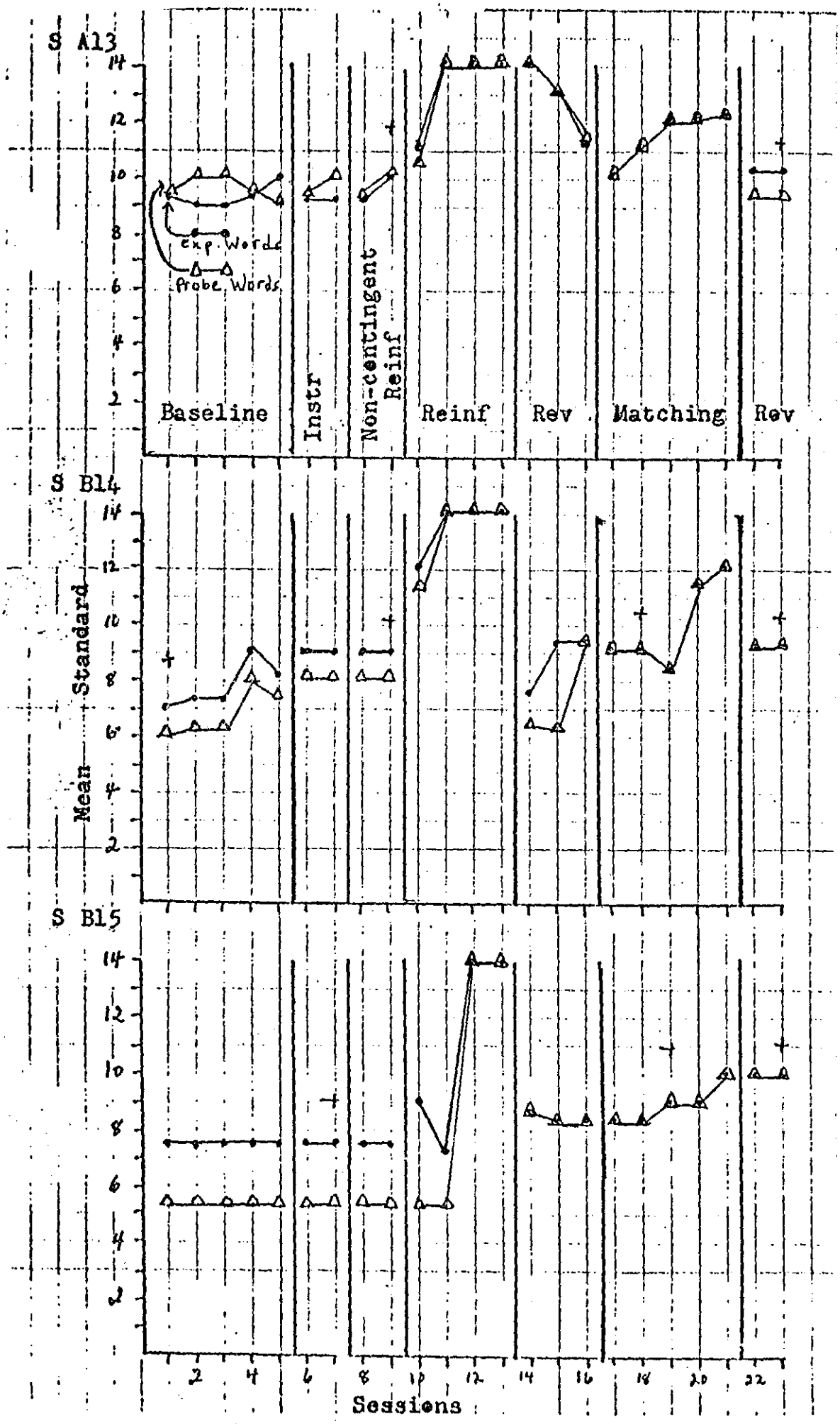
Figure 2. Mean standard for each subject (+ = 1 exp. session).

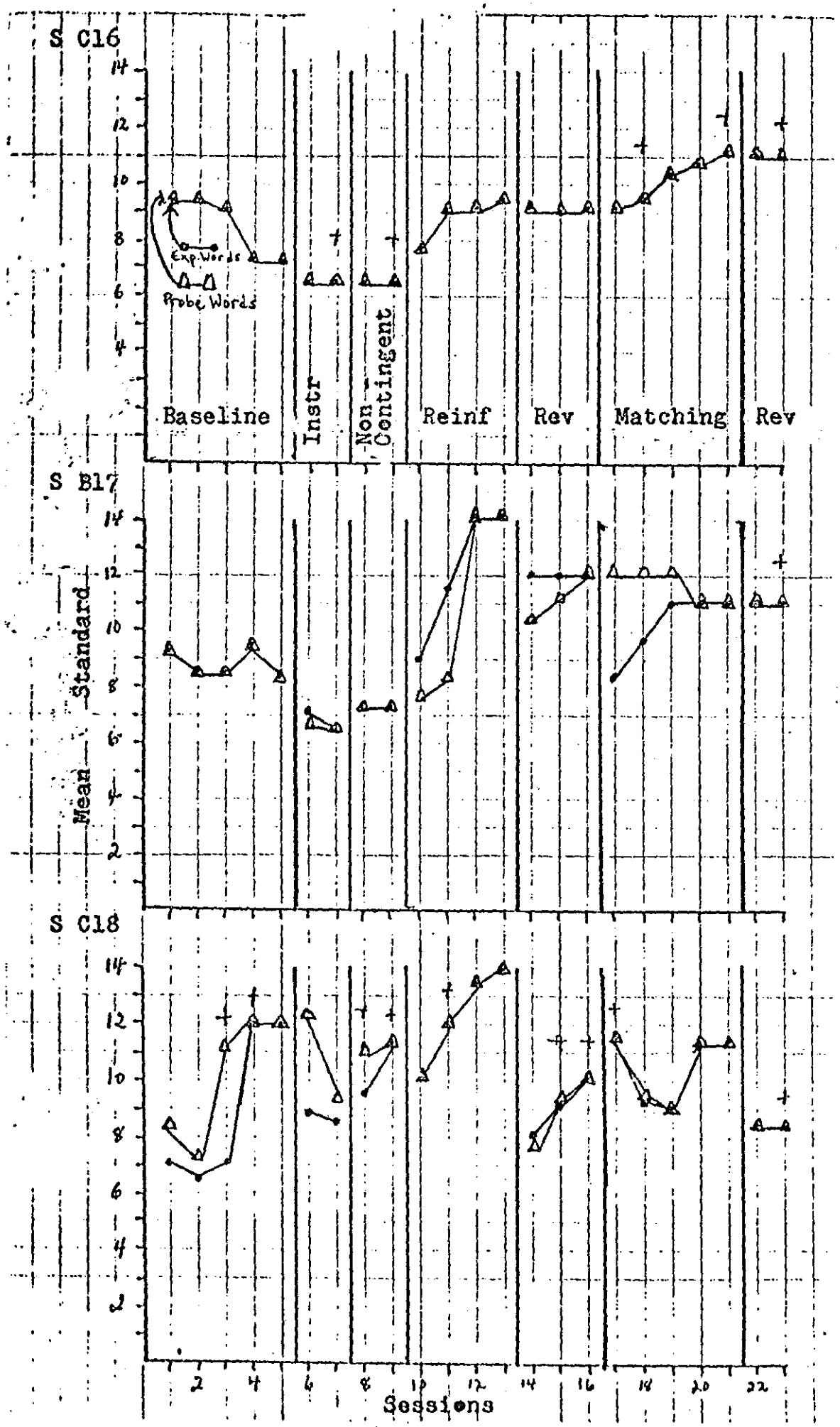


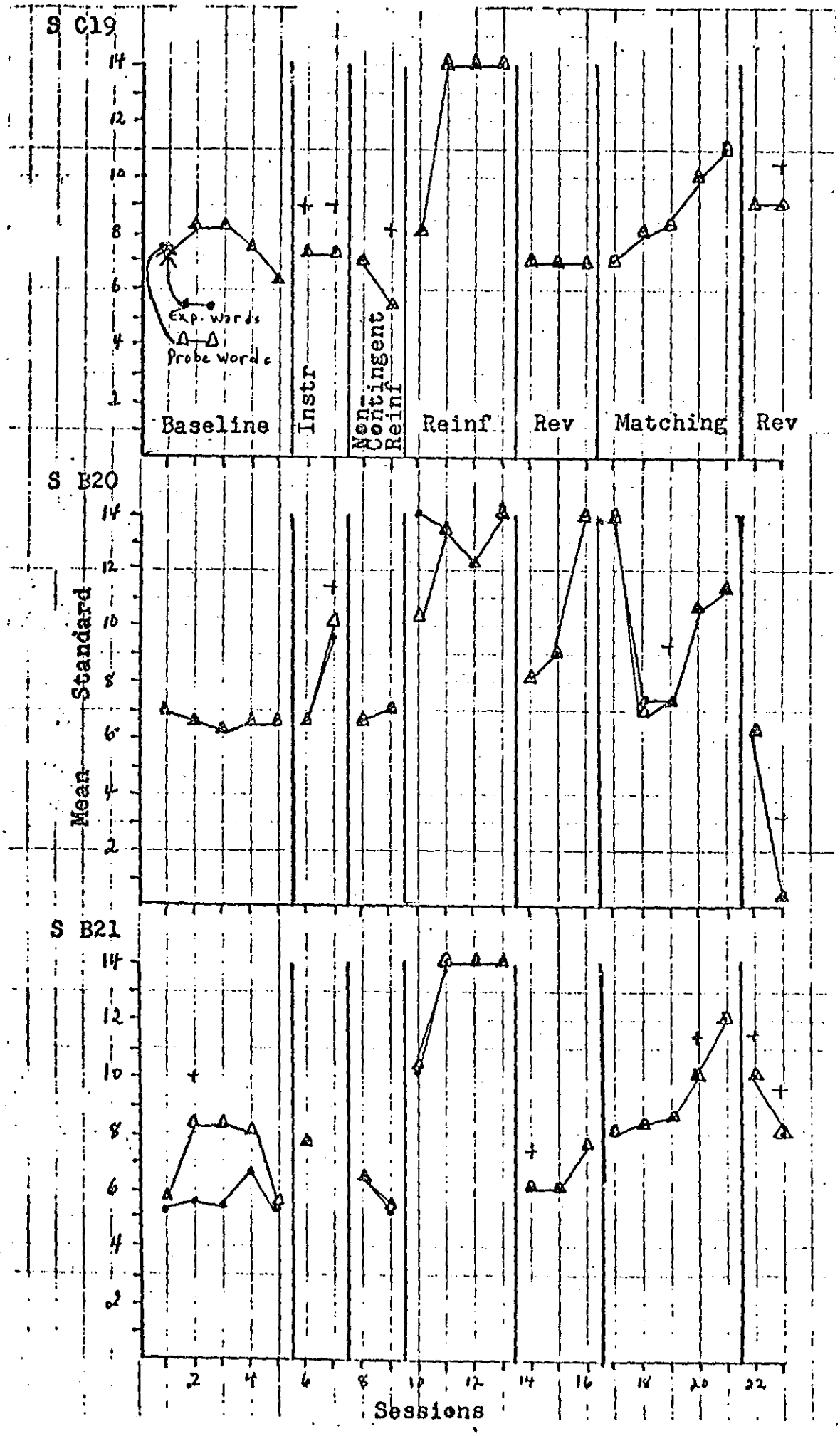

S A4
S A5
S A6
Mean
Standard
Exp Words
Probe Words
Baseline
Instr
Non-Contingent Reinf
Reinf
Rev
Matching
Rev
Sessions

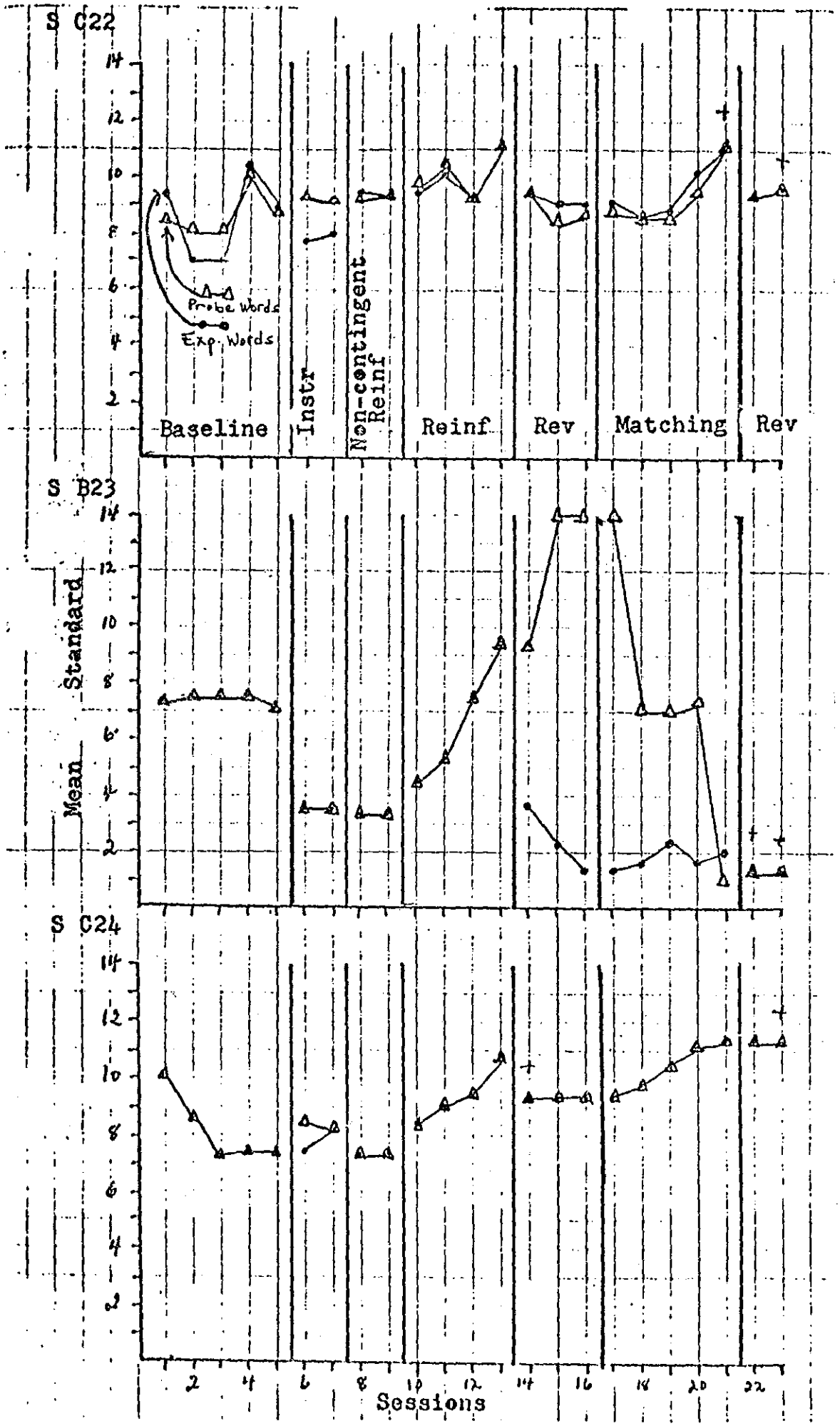

S A7
S A8
S A9
Mean Standard
Probe Words
Exp. Words
Baseline
Instr
Non-contingent Reinf
Reinf
Rev
Matching
Rev
Sessions

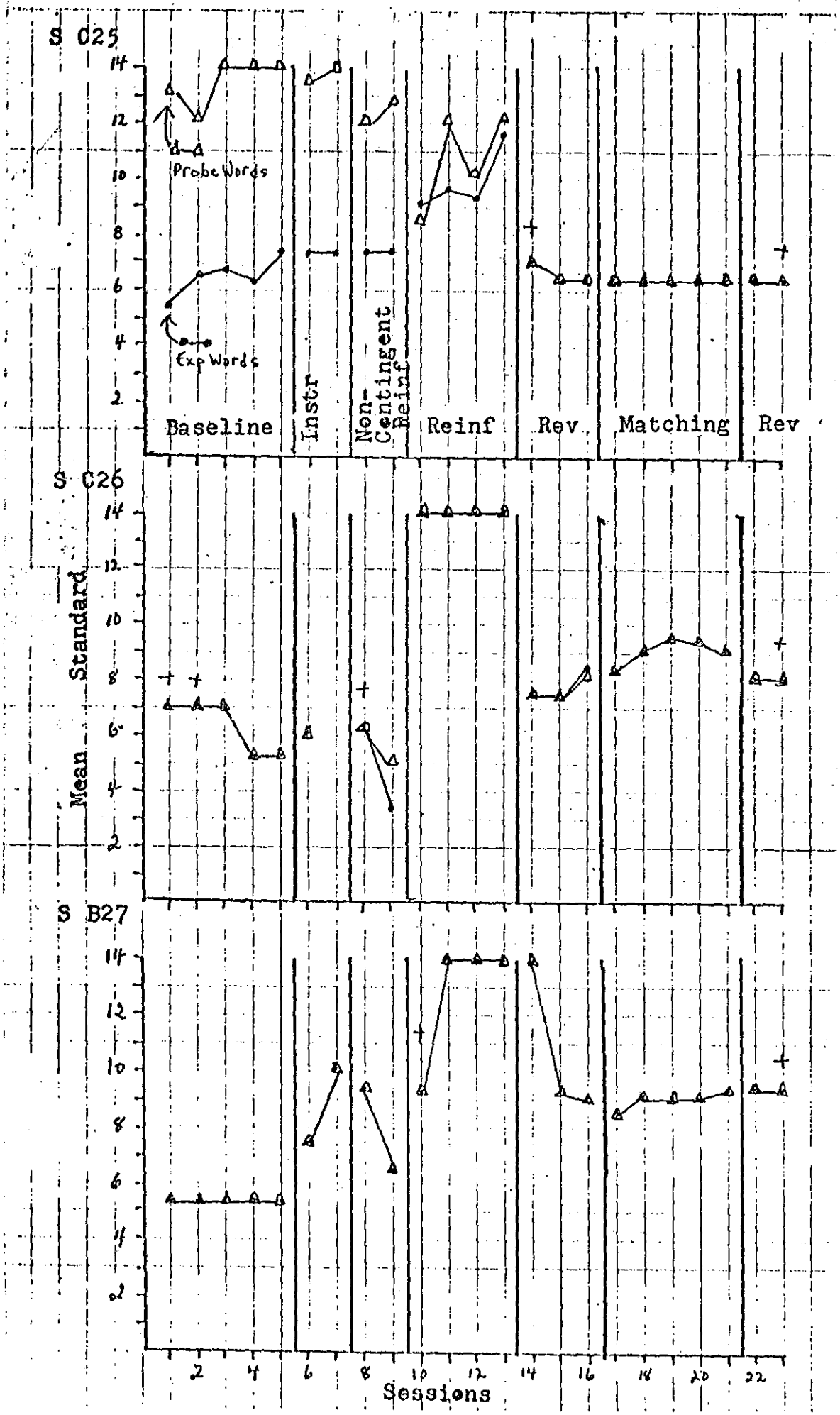

S A10
S A11
S A12
Mean Standard
Exp. Words
Probe Words
Baseline
Instr
Non-contingent Reinf
Reinf
Rev
Matching
Rev
Sessions

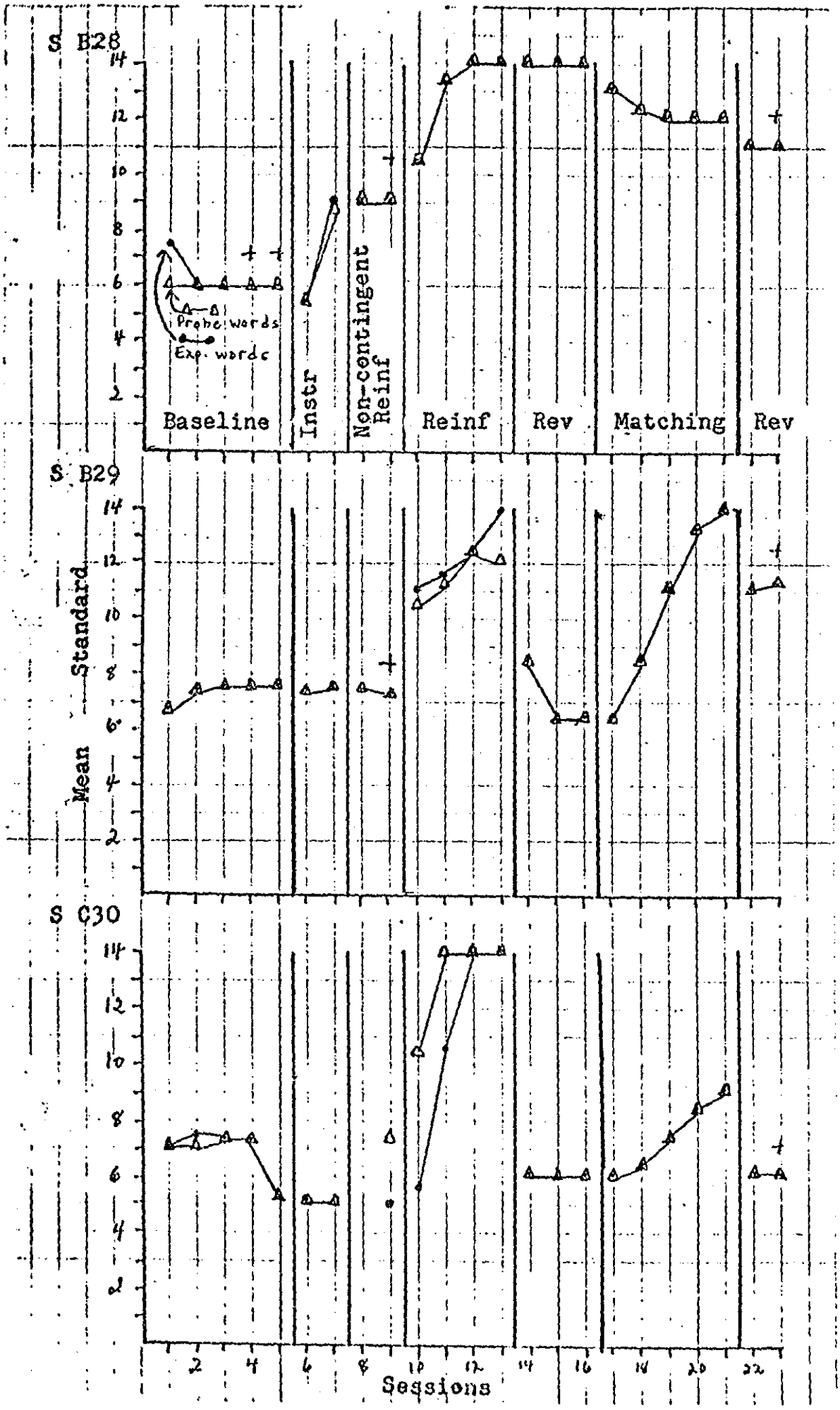

S A13
S B14
S B15
Mean Standard
Exp. Words
Probe Words
Baseline
Instr
Non-contingent Reinf
Reinf
Rev
Matching
Rev
Sessions


S C16
S B17
S C18
Mean Standard
Exp. Words
Probe Words
Baseline
Instr
Non Contingent
Reinf
Rev
Matching
Rev
Sessions


S C19
S B20
S B21
Mean Standard
Exp. words
Probe words
Baseline
Instr
Non-Contingent Reinf
Reinf
Rev
Matching
Rev
Sessions


S C22
S B23
S C24
Mean Standard
Probe Words
Exp. Words
Baseline
Instr
Non-contingent Reinf
Reinf
Rev
Matching
Rev
Sessions


S C25
Probe Words
Exp Words
Baseline
Instr
Non-Contingent Reinf
Reinf
Rev
Matching
Rev
S C26
S B27
Mean Standard
Sessions


S B28
S B29
S C30
Mean Standard
Probe words
Exp. words
Baseline
Instr
Non-contingent Reinf
Reinf
Rev
Matching
Rev
Sessions

single subject graphs were ordered such that subjects setting higher standards during baseline appear first.

Baseline. For comparison purposes, subjects were classified into a high standard group (Group A), a low standard group (Group B), and a variable standard group (Group C), based on the level of self-determined standards they set for baseline. The group of which a subject is a member can be determined by the letter (A, B, or C) preceding his/her subject number on the graphs in Figure 2 as well as within the text itself.

The high standard group, Group A, included Subjects 1-13. These subjects' mean standard level ranged from 12-9, and remained fairly constant throughout baseline. Subjects 14, 15, 17, 20, 21, 23, 27, 28, 29, formed the low standard group (Group B) with mean standards ranging from 8.66-5.33. The third group (Group C) included Subjects 16, 18, 19, 22, 24, 25, 26, 30, who exhibited the greatest changes in their mean standard level during baseline. Subjects 16, 24, 26, 30 gradually lowered their mean standard level, while Subjects 18 and 25 raised their mean standard. Subject 22 increased and decreased his mean standard throughout baseline. The mean standard level for this group ranged from 12-5.

Instructions. Twenty-two subjects demonstrated little if any change in the level of their mean standard during the four day instruction phase. Three subjects, Ss B17, B23, and C18, lowered their mean standard by two or more points, while five subjects, Ss A5, B20, B21,

B27, and B28 raised their mean standard by approximately the same amount.

Non-contingent reinforcement. Twenty-three subjects maintained standards at approximately the same level during this condition. Six subjects, Ss A9, B20, B21, B27, C19, and C26 lowered standards by two or more points. One subject, S C18, raised her mean standard by 3 points.

Contingent reinforcement for raising standards. In this phase all subjects raised their mean standard level with 25 subjects setting standards at the maximum level (14) by the end of the phase. Four subjects, Ss B23, C16, C24, and C25, raised their mean standard by 4-6 points but did not reach maximum. Subject C22 demonstrated the most variable standard setting behavior of all the subjects. His mean standard at the end of the phase exceeded its initial level by 2 points.

Reversal I. Seventeen subjects lowered standards to approximately baseline levels in this phase. These subjects were Ss A2, A5, A6, A13, B14, B15, B21, B23, B27, B29, C18, C19, C22, C24, C25, C26, and C30. Eleven subjects did not lower standards at all (Ss A1, A4, A8, A11, A12, B28, C16) or retained a mean standard above baseline (Ss A3, A7, A10, B17). Two subjects, S A9 and S B20, decreased their mean standard and then increased it to near treatment levels.

Matching standards and performance. Fifteen subjects raised standards during the matching condition with the magnitude of increase varying for each subject and ranging from 1-7 points. These subjects were Ss A2, A5, A6, A13, B14, B15, B29, C16, C18, C19, C22, C24, C26, and C30. Nine subjects (Ss A1, A3, A4, A7, A8, A9, A10, A11, A12) whose mean standard was close to or at the maximum level at the onset of this

condition essentially maintained standards at this level. Five of these subjects (Ss A1, A3, A7, A9, A11) however exhibited some decreases and then increases in their mean standard level during this phase but reached maximum within 1 point by the end of the condition. Three subjects, Ss B17, B20, and B28, demonstrated an overall decrease in their mean standard level although Ss B17 and B20, began increasing their standards again in the latter part of the matching phase. Three subjects, Ss B23, B27, and C25, demonstrated little change in their mean standard for the duration of this condition.

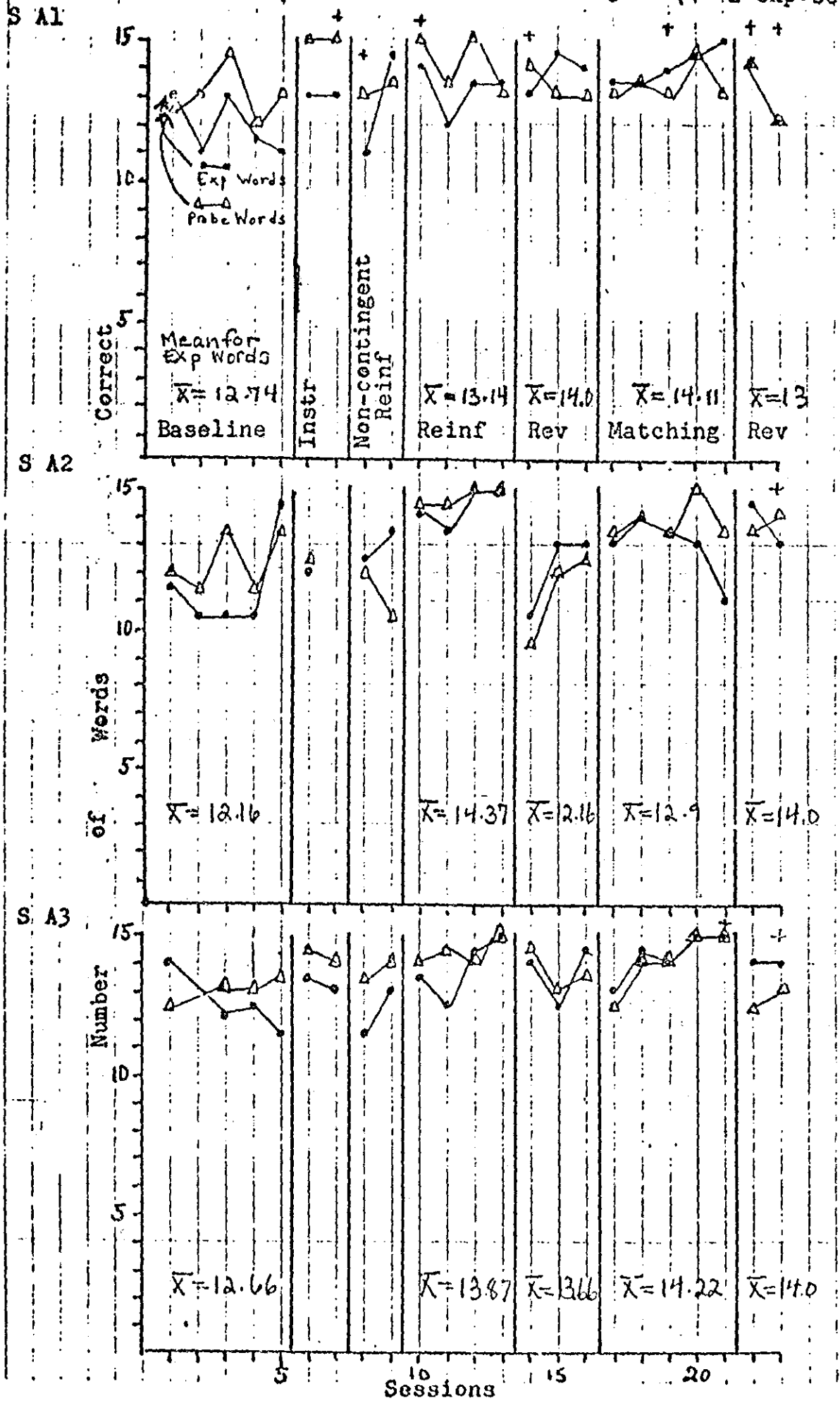
Reversal II. Seven subjects, Ss A13, B14, B20, B21, B22, C18, and C30, lowered standards to approximately baseline levels, representing decreases of varying magnitudes. Nineteen subjects either maintained standards at the same level as during the matching condition (Ss A4, A5, A6, A8, A9, A10, A12, B15, B17, C16, C24), or maintained standards at a level above baseline (Ss A1, A2, A3, A7, A11, B29, C19, C26). Eight of these 19 subjects had lowered standards to baseline levels in the preceding reversal I phase; nine of them were setting standards around the maximum level (14) since the contingent reinforcement phase; and two of them exhibited changes in their level of self-determined standards only during the reinforcement phases. The three subjects (Ss B23, C25, C27) who had demonstrated little change in their mean standard during matching maintained standards at the same level during this reversal phase, while S B28 continued to slightly lower standards in this phase.

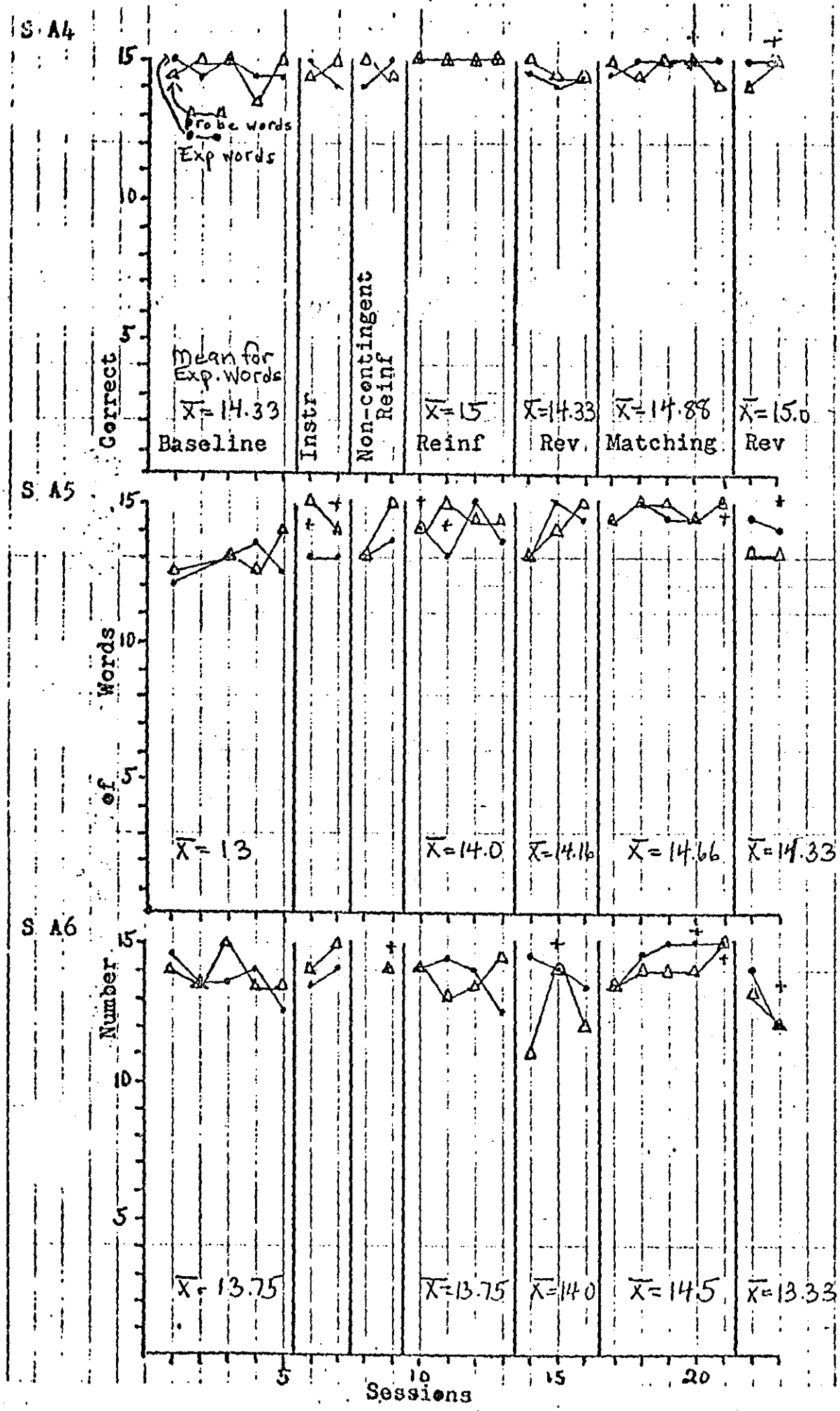
Although subjects did not always set standards for the probe words at the same level as the standards they set for the experimental words, changes in the mean standard level were in the same direction for both sets of words. Standards for probe words were particularly close to standards for experimental words during the contingent reinforcement and matching phases.

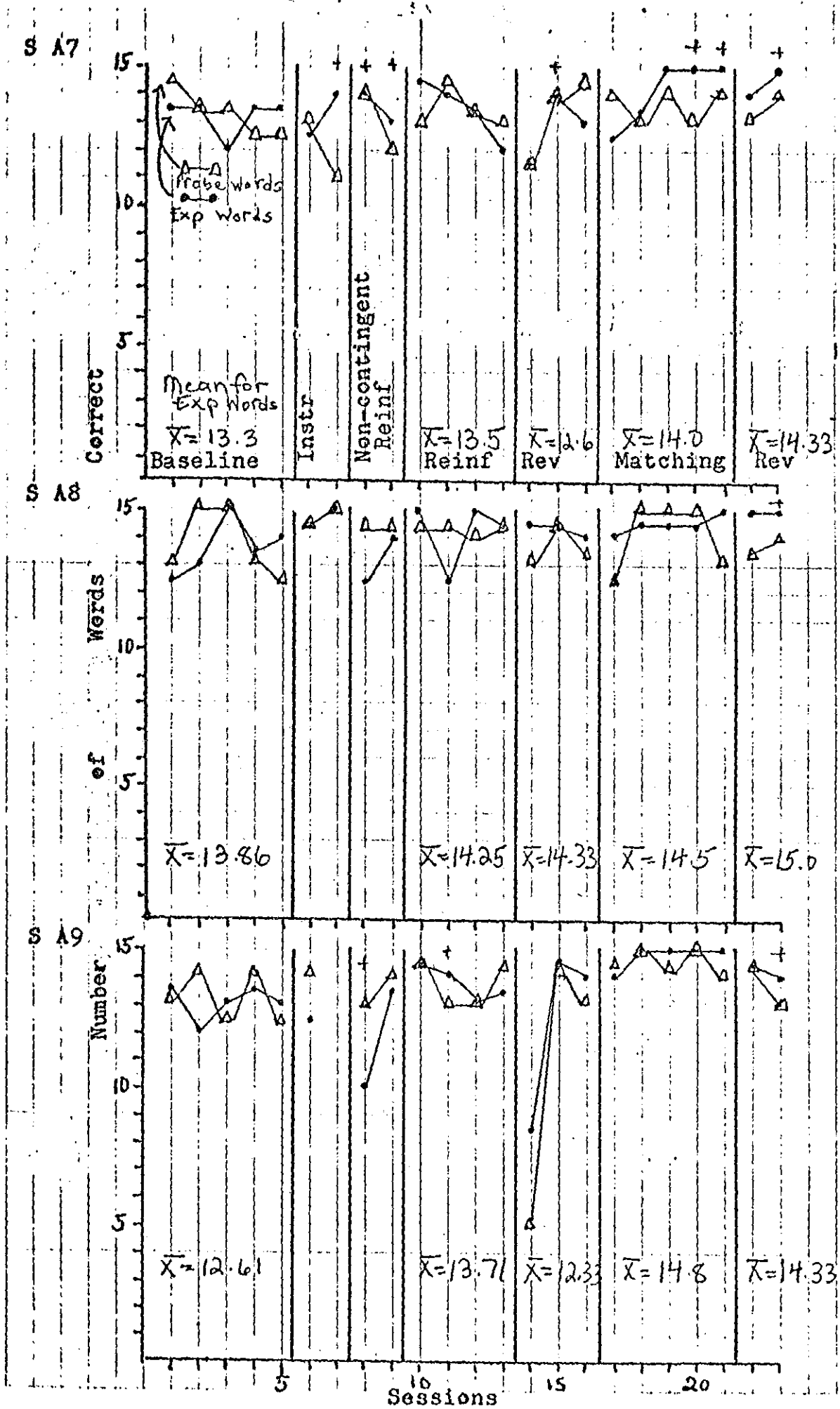
Performance. Figure 3 represents the number of words spelled correctly (level of performance) by each subject for both the experimental and probe words, across all conditions. Data were combined across days such that each data point represents two experimental sessions. A plus mark (+) above a particular data point signifies the days upon which a data point represents a single experimental session. The single-subject graphs were ordered so that they correspond to the single-subject graphs for standards. In other words S 1's graphs appear first in both Figure 2 and Figure 3 and so on for the other 29 subjects.

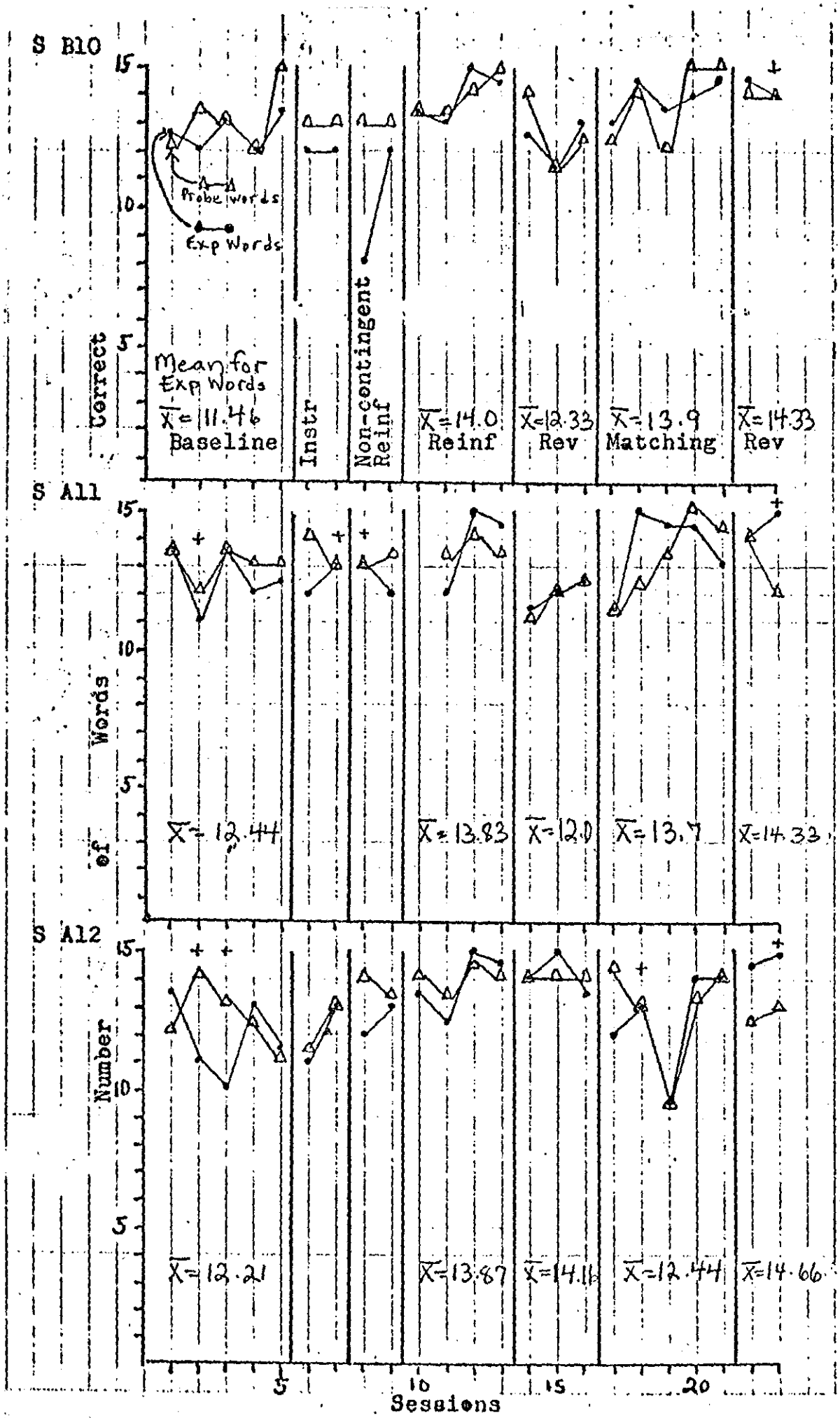
Each subject's mean performance for a particular phase was used to assess changes in the level of performance. Performance scores for the first three phases, baseline, instructions, and non-contingent reinforcement, were combined to form one measure of mean performance for baseline for each subject. This was done because the experimental manipulations during the instructions and non-contingent reinforcement phases were applied to the level of self-determined standards. Since approximately 80% of the subjects did not change their level of self-determined standards during these two phases, any changes that did occur, probably, cannot be attributed to the experimental manipulations employed. Hence

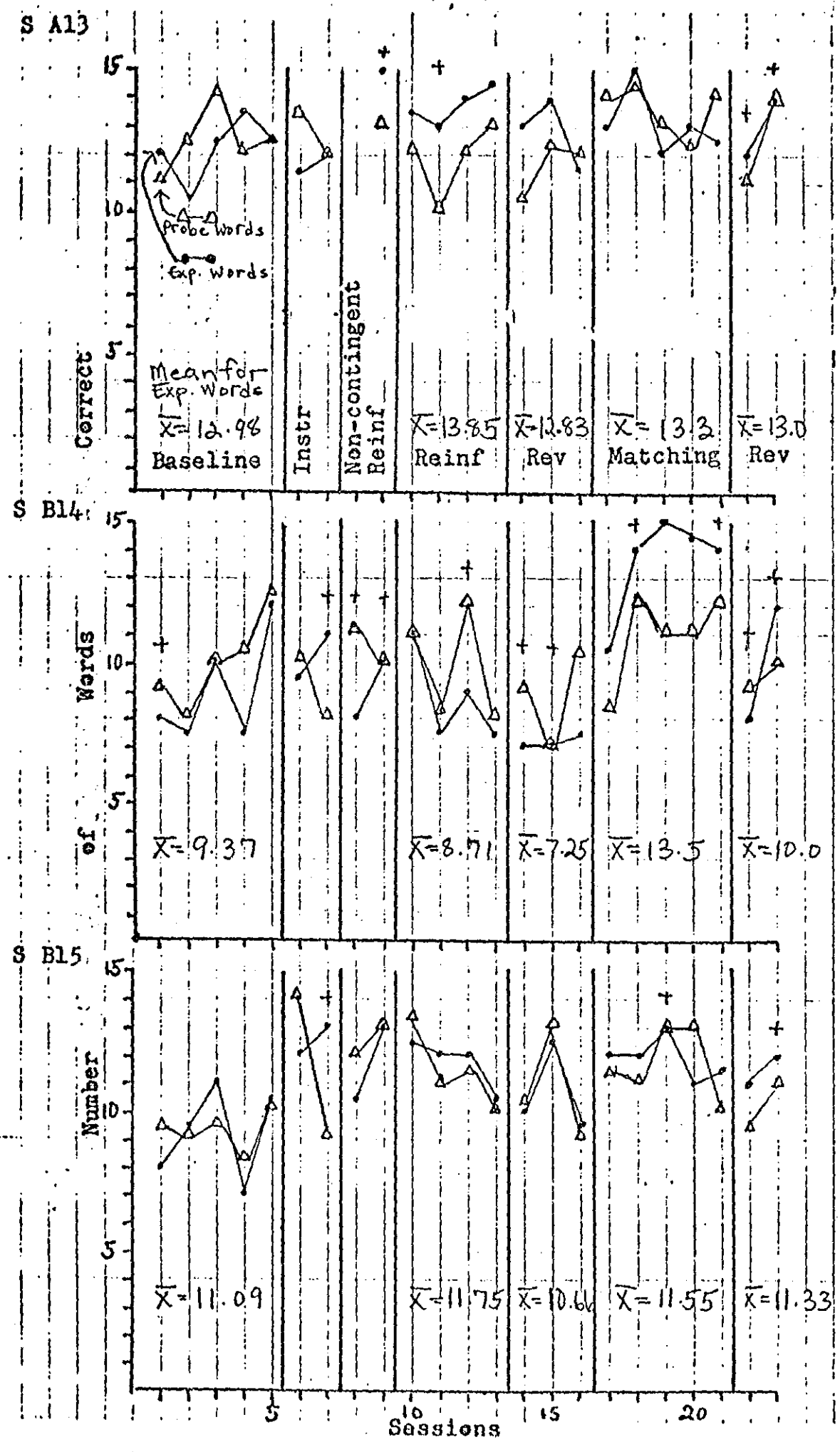
Level of performance for each subject (+ = 1 exp. session)

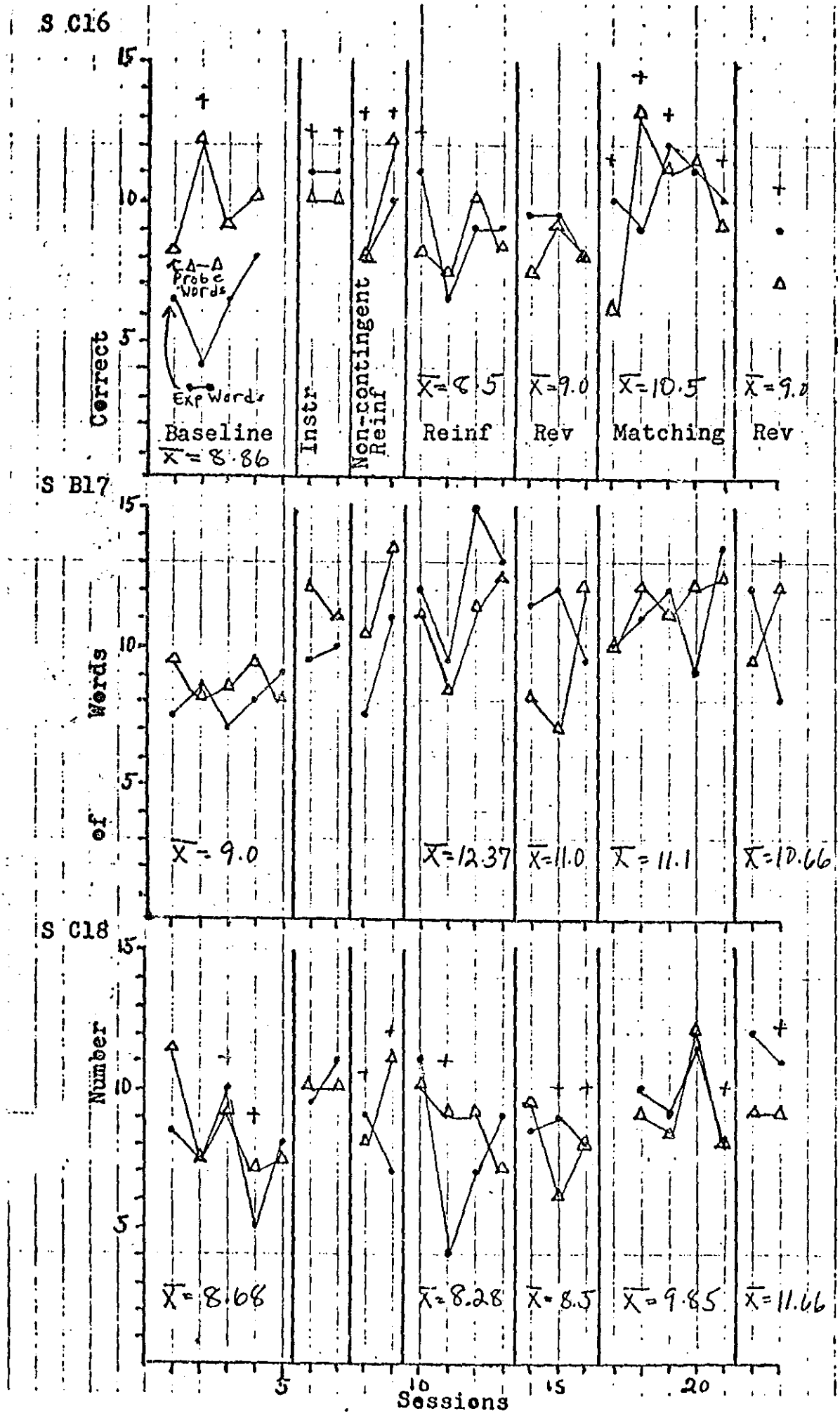


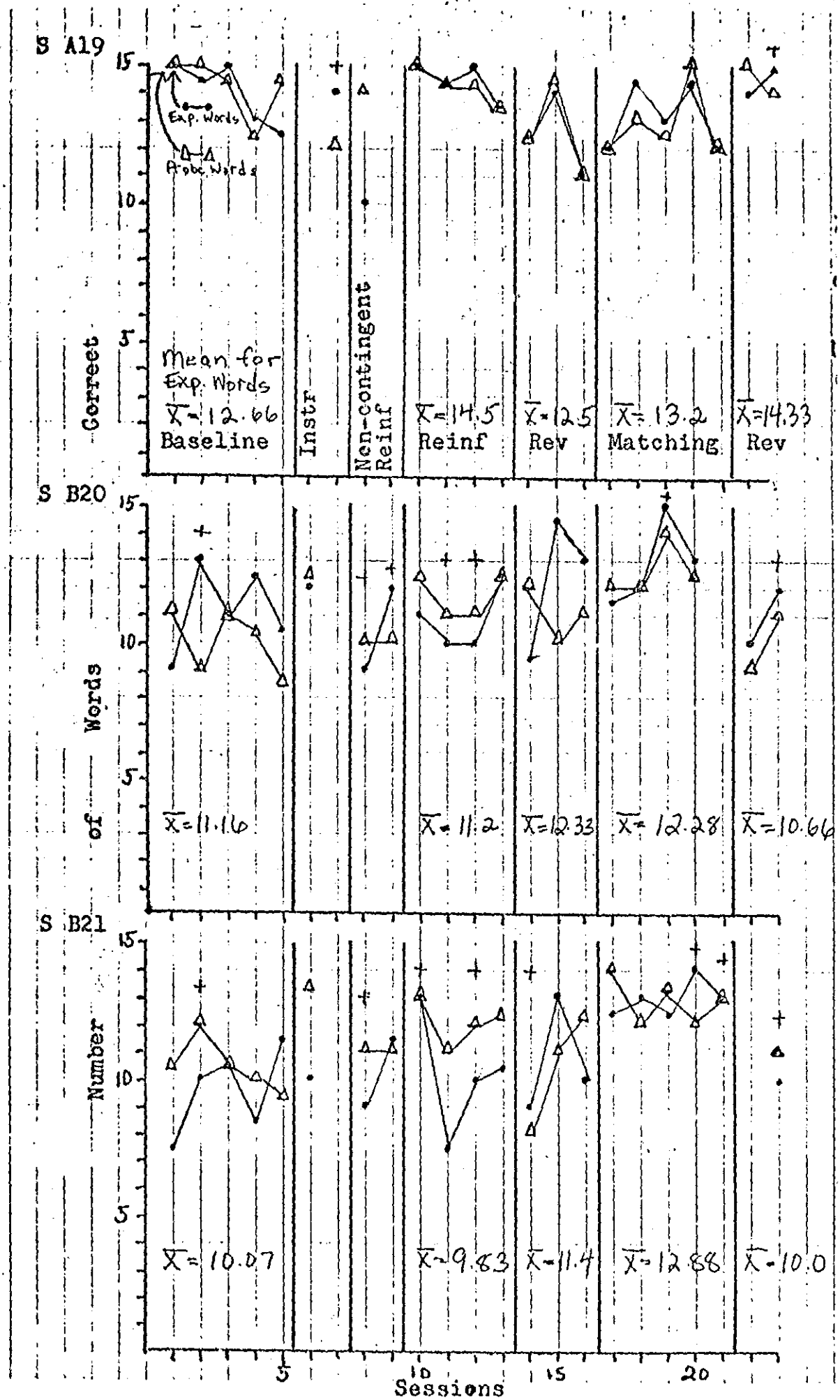

S. A4
S. A5
S. A6
Correct
Words
of
Number
Probe words
Exp words
Mean for Exp. Words
X̄ = 14.33
Baseline
Instr
Non-contingent Reinf
X̄ = 15
Reinf
X̄ = 14.33
Rev.
X̄ = 14.88
Matching
X̄ = 15.0
Rev
X̄ = 13
X̄ = 14.0
X̄ = 14.16
X̄ = 14.66
X̄ = 14.33
X̄ = 13.75
X̄ = 13.75
X̄ = 14.0
X̄ = 14.5
X̄ = 13.33
Sessions

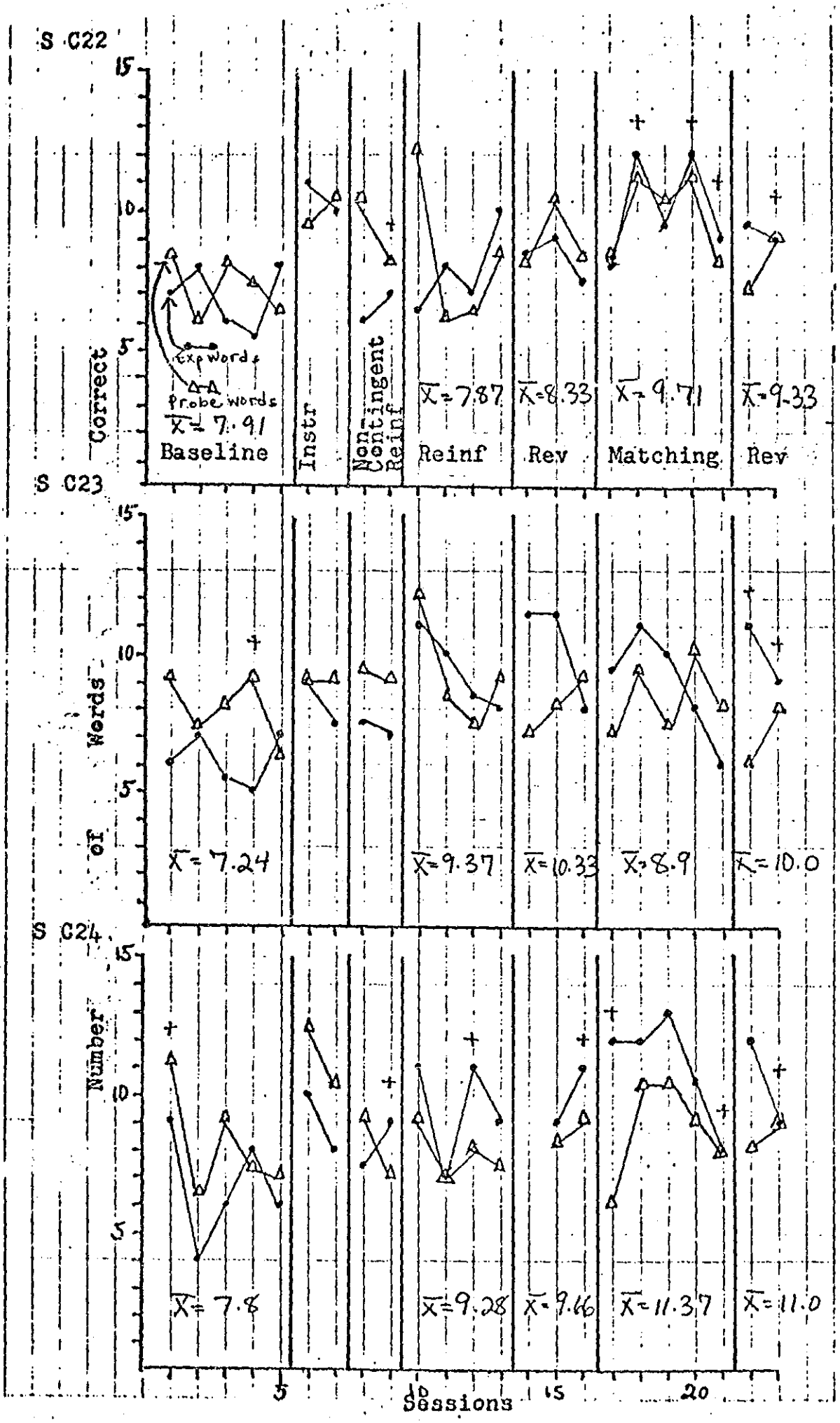

S A7
15
10
5
Correct
Probe words
Exp Words
Mean for Exp Words
X̄ = 13.3
Baseline
Instr
Non-contingent Reinf
X̄ = 13.5
Reinf
X̄ = 12.6
Rev
X̄ = 14.0
Matching
X̄ = 14.33
Rev
S A8
15
10
5
Words of
X̄ = 13.86
X̄ = 14.25
X̄ = 14.33
X̄ = 14.5
X̄ = 15.0
S A9
15
10
5
Number
X̄ = 12.61
X̄ = 13.71
X̄ = 12.33
X̄ = 14.8
X̄ = 14.33
5
10
15
20
Sessions

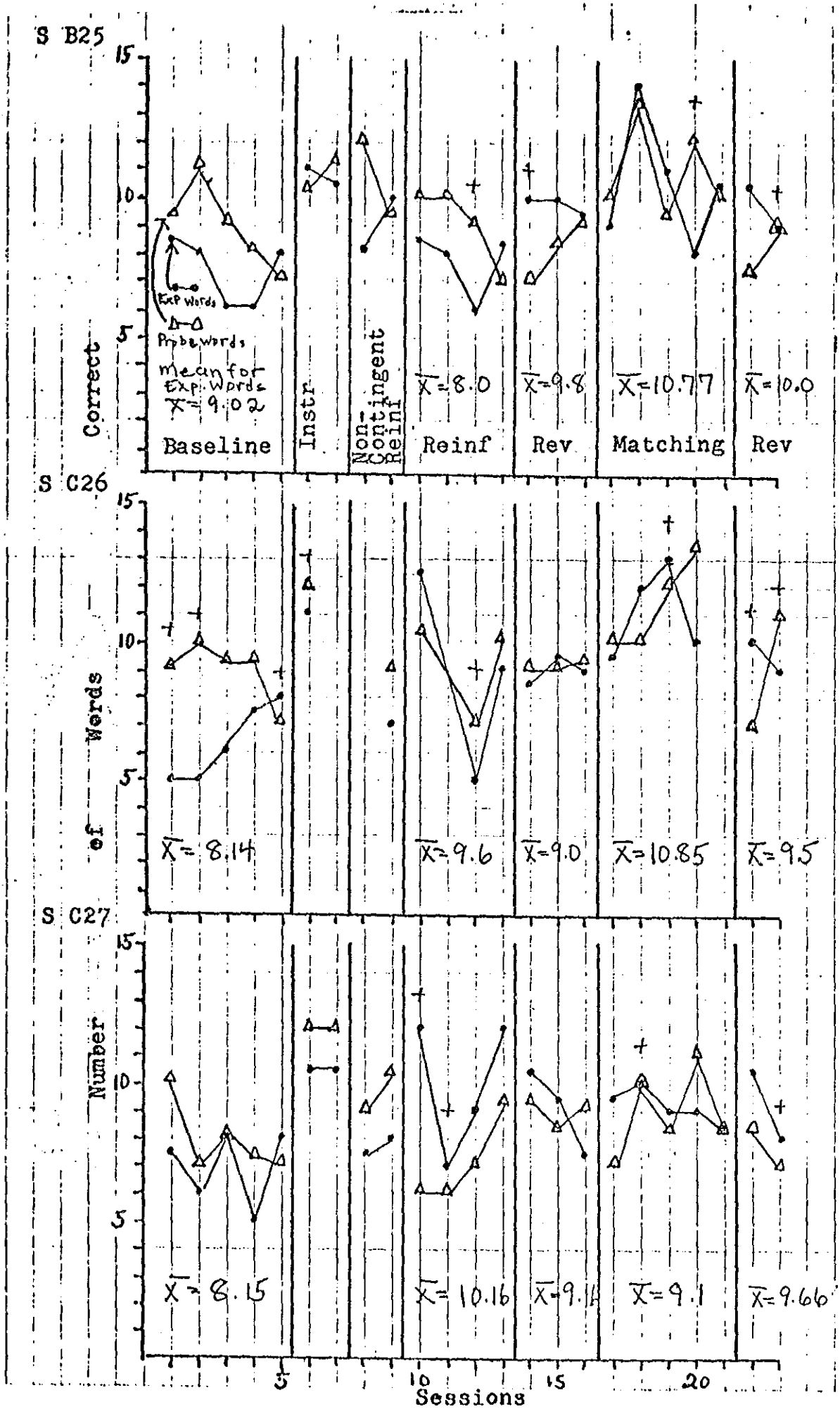

S B10
15
10
5
Probe words
Exp Words
Mean for Exp Words
$\bar{X}$ = 11.46
Baseline
Instr
Non-contingent Reinf
$\bar{X}$ = 14.0
Reinf
$\bar{X}$ = 12.33
Rev
$\bar{X}$ = 13.9
Matching
$\bar{X}$ = 14.33
Rev
Correct
S A11
$\bar{X}$ = 12.44
$\bar{X}$ = 13.83
$\bar{X}$ = 12.0
$\bar{X}$ = 13.7
$\bar{X}$ = 14.33
Words
of
S A12
$\bar{X}$ = 12.21
$\bar{X}$ = 13.87
$\bar{X}$ = 14.16
$\bar{X}$ = 12.44
$\bar{X}$ = 14.66
Number
5
10
15
20
Sessions

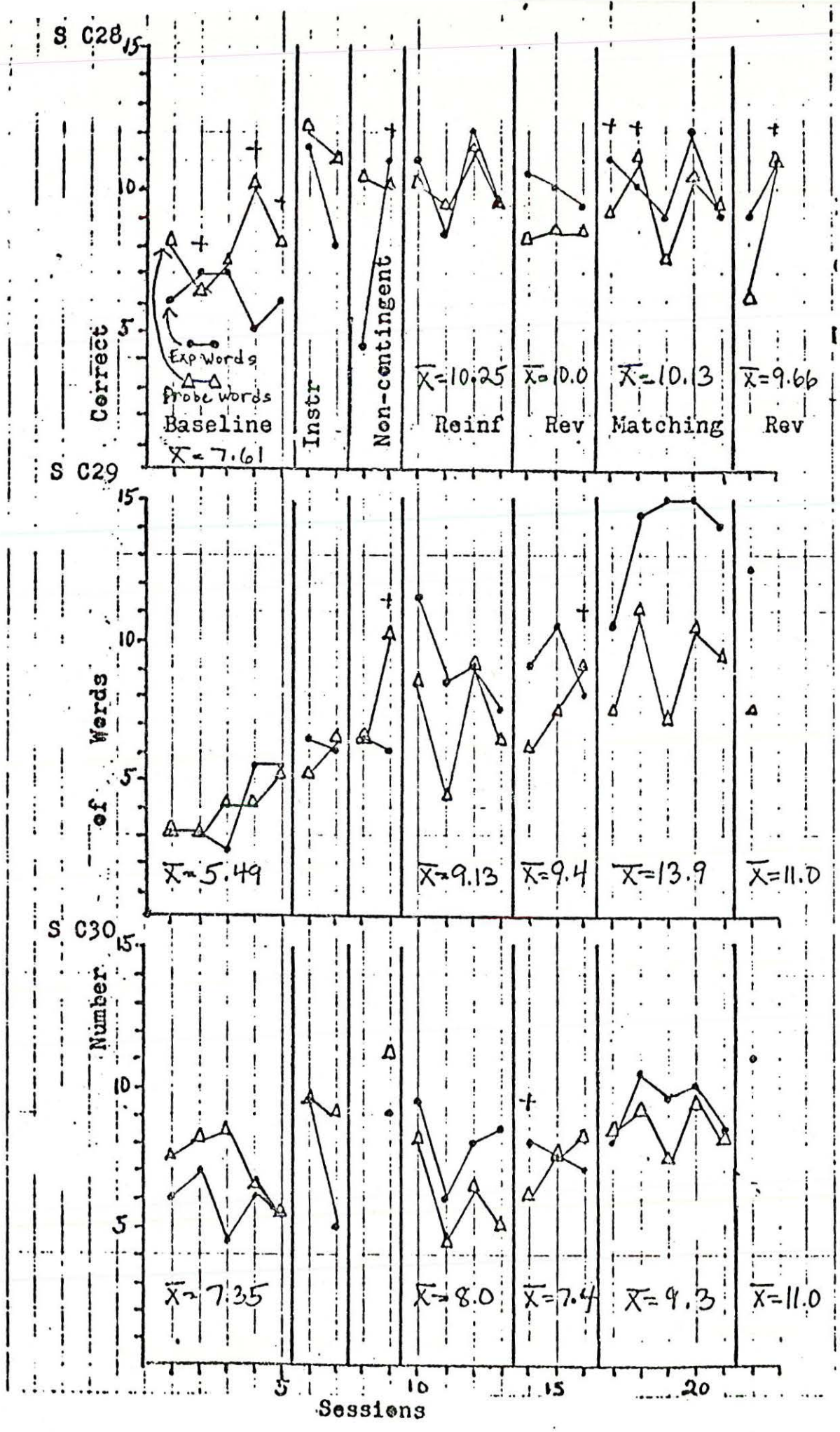

S A13
15
10
5
Correct
Probe Words
Exp. Words
Mean for Exp. Words
X̄=12.98
Baseline
Instr
Non-contingent Reinf
X̄=13.85
Reinf
X̄=12.83
Rev
X̄=13.2
Matching
X̄=13.0
Rev
S B14
Words
X̄=9.37
X̄=8.71
X̄=7.25
X̄=13.5
X̄=10.0
S B15
Number
X̄=11.09
X̄=11.75
X̄=10.66
X̄=11.55
X̄=11.33
5
10
15
20
Sessions


S C16
Correct
15
10
5
Probe Words
Exp Words
Baseline
X̄ = 8.86
Instr
Non-contingent Reinf
X̄ = 9.5
Reinf
X̄ = 9.0
Rev
X̄ = 10.5
Matching
X̄ = 9.0
Rev
S B17
Words
of
15
10
5
X̄ = 9.0
X̄ = 12.37
X̄ = 11.0
X̄ = 11.1
X̄ = 10.66
S C18
Number
15
10
5
X̄ = 8.68
X̄ = 8.28
X̄ = 8.5
X̄ = 9.85
X̄ = 11.66
5
10
15
20
Sessions


S A19
Exp. Words
Probe Words
Correct
Mean for Exp. Words
$\bar{X}$ = 12.66
Baseline
Instr
Non-contingent Reinf
$\bar{X}$ = 14.5
Reinf
$\bar{X}$ = 12.5
Rev
$\bar{X}$ = 13.2
Matching
$\bar{X}$ = 14.33
Rev
S B20
Number of Words
$\bar{X}$ = 11.16
$\bar{X}$ = 11.2
$\bar{X}$ = 12.33
$\bar{X}$ = 12.28
$\bar{X}$ = 10.66
S B21
$\bar{X}$ = 10.07
$\bar{X}$ = 9.83
$\bar{X}$ = 11.4
$\bar{X}$ = 12.88
$\bar{X}$ = 10.0
15
10
5
5
10
15
20
Sessions


S C22
15
10
5
Correct
exp words
probe words
X̄ = 7.91
Baseline
Instr
Non-Contingent Reinf
X̄ = 7.87
Reinf
X̄ = 8.33
Rev
X̄ = 9.71
Matching
X̄ = 9.33
Rev
S C23
15
10
5
of Words
X̄ = 7.24
X̄ = 9.37
X̄ = 10.33
X̄ = 8.9
X̄ = 10.0
S C24
15
10
5
Number
X̄ = 7.8
X̄ = 9.28
X̄ = 9.66
X̄ = 11.37
X̄ = 11.0
5
10
15
20
Sessions


S B25
S C26
S C27
Correct
Number of Words
Exp Words
Probe words
mean for Exp. Words X̄ = 9.02
Baseline
Instr
Non-contingent Reinf
Reinf
Rev
Matching
Rev
X̄ = 8.0
X̄ = 9.8
X̄ = 10.77
X̄ = 10.0
X̄ = 8.14
X̄ = 9.6
X̄ = 9.0
X̄ = 10.85
X̄ = 9.5
X̄ = 8.15
X̄ = 10.16
X̄ = 9.1
X̄ = 9.66
15
10
5
5
10
15
20
Sessions


S C28
Correct
15
10
5
Exp words
Probe words
Baseline
X̄=7.61
Instr
Non-contingent
X̄=10.25
Reinf
X̄=10.0
Rev
X̄=10.3
Matching
X̄=9.66
Rev
S C29
Words
of
15
10
5
X̄=5.49
X̄=9.13
X̄=9.4
X̄=13.9
X̄=11.0
S C30
Number
15
10
5
X̄=7.35
X̄=8.0
X̄=7.4
X̄=9.3
X̄=11.0
5
10
15
20
Sessions

performance scores obtained during these two conditions represent a continuation of baseline.

Baseline. For comparison purposes subjects were classified into a high scoring group (Group A), a middle scoring group (Group B), and a low scoring group (Group C), based on their mean performance scores for baseline. The group of which a subject is a member can be determined by the letter (A, B, or C) preceding his/her subject number on the graphs in Figure 3, as well as within the text itself.

The high scoring group, Group A, included Subjects 1-9, 11, 12, 13, and 19. Mean performance scores for these subjects ranged from 12-14.33. Subjects 10, 14, 15, 17, 20, 21, and 25 scored in the middle range (9-11), forming Group B. The third group, Group C, included Subjects 16, 18, 22, 23, 24, and 26-30. These subjects obtained a mean performance score ranging from 5.5-8.

Contingent reinforcement for raising standards. Twenty subjects obtained mean performance scores of 1/2 point or more above their mean performance scores for baseline. Twelve of these subjects, Ss A1, A3, A4, A5, A9, A11, A12, A13, B15, C24, C26, and C30, improved their mean performance score by .5-1.6 points. The other eight subject, Ss A2, A19, B10, B17, C23, C27, C28, and C29, improved their mean performance scores by 2 or more points. Eight subjects, Ss A6, A7, A8, B20, B21, C16, C18, and C22, maintained the same mean score as in baseline. Two subjects, Ss B14, and B25, obtained a lower mean performance score by .5 and 1 point, respectively.

Reversal I. The mean performance score for nine subjects (Ss A1, A2, A4, A9, A11, A13, A19, B10, B15, C30) decreased to near baseline level in this phase. Three subjects, Ss B17, C26, and C27, obtained a lower mean performance score by .6-1.4 points, but still obtained a score above baseline. Two subjects, S A7 and S B14, obtained a lower mean performance score by .9 and 1.5 points respectively. Neither of these two subjects had improved performance in the previous condition. Nine subjects, Ss A3, A5, A6, A8, A12, C18, C24, C28, and C29, obtained the same mean performance score as in the previous phase, and seven subjects, Ss A1, B20, B21, C16, C22, C23, and C25, obtained a higher mean performance score by .5-1.8 points.

Matching standards and performance. Twenty-two subjects obtained a higher mean performance score in this phase by 1/2 point or more as compared to the score they obtained in the reversal I phase. Fourteen of these subjects (Ss A2, A3, A4, A5, A6, A7, A13, A19, B15, B21, B25, C16, C18, C22) improved their mean performance score by .5-1.5 points. The other eight subjects (Ss A9, A11, B10, B14, C24, C26, C29, C30) improved their mean performance scores by more than 1.5 points, with S B14 and S C29 demonstrating the largest increases of 6.5 points and 4.5 points respectively. In addition, 16 of the 22 subjects who had improved performance during this phase obtained mean performance scores that exceeded their mean performance scores during the contingent reinforcement phase. The other six subjects who improved performance in this phase, obtained mean performance scores equal to or slightly lower than their mean performance scores in the contingent reinforcement phase.

Six subjects, Ss A1, A8, B17, B20, C27, and C28, obtained the same mean performance score as in the previous phase. Two subjects, S A12, and S B23, obtained a lower mean score by 1.7 and 1.4 points respectively.

Reversal II. Two subjects, S B25 and X C16, obtained the same mean performance score as they had obtained in the reversal I phase, reflecting decreases of .7 and 1.5 points. Four subjects, Ss A9, B14, C26, and C29, obtained a lower mean performance score, but remained at a level above their mean performance score during the reversal I phase. Two subjects, S A6 and S B21, who had been attaining slightly higher mean performance scores since baseline, obtained lower mean performance scores in this phase by 1 point and 2.8 points respectively. Four subjects, Ss A1, B17, B20, and C28, whose performance scores had remained relatively unchanged during the matching phase obtained lower mean performance scores by .5-1.5 points. Eighteen subjects either obtained the same mean performance score in this phase (Ss A3, A4, A5, A7, A13, B10, B15, C22, C24), or obtained a higher mean performance score by .5 to 2 points (Ss A2, A8, A11, A12, A19, C18, C23, C27, C30).

Mean performance scores for the probe words changed in the same direction as mean performance scores for the experimental words although the magnitude of change often differed for the two scores. Subjects A13 and C29 exhibited smaller changes in their mean performance scores for the probe words during the contingent reinforcement phase in one case and during the matching condition in the other.

Group Correlations Between the Level of Self-Determined Standards and Performance

Figure 4 presents the Pearson r linear correlation between the mean standard and the level of performance for the experimental words, for all subjects, across all conditions. Data were combined across days such that each data point represents two experimental sessions. A plus mark (+) above a particular data point indicates the days upon which a data point represents a single experimental session.

The mean correlation coefficient between the mean standard and performance during baseline equalled .55, with a range of .50-.64. The mean correlation coefficient for the instructions and non-contingent reinforcement phases equalled .52 and .65 respectively. The mean correlation coefficient was slightly lower for the contingent reinforcement condition and equalled .42, ranging from .34-.50. During the reversal I phase the mean correlation coefficient increased to .60 and increased further in the matching phase to .65. The linear correlation index was especially high for the last two days of the matching phase (.82 and .77). In the second reversal phase the mean correlation coefficient returned to its former level of .60.

Discussion

The results from this study confirm earlier reports suggesting that students will set low standards for themselves in order to maximize reinforcement (Felixbrod & O'Leary, 1972). With respect to the academic task for which subjects set standards in this study, a standard of 11

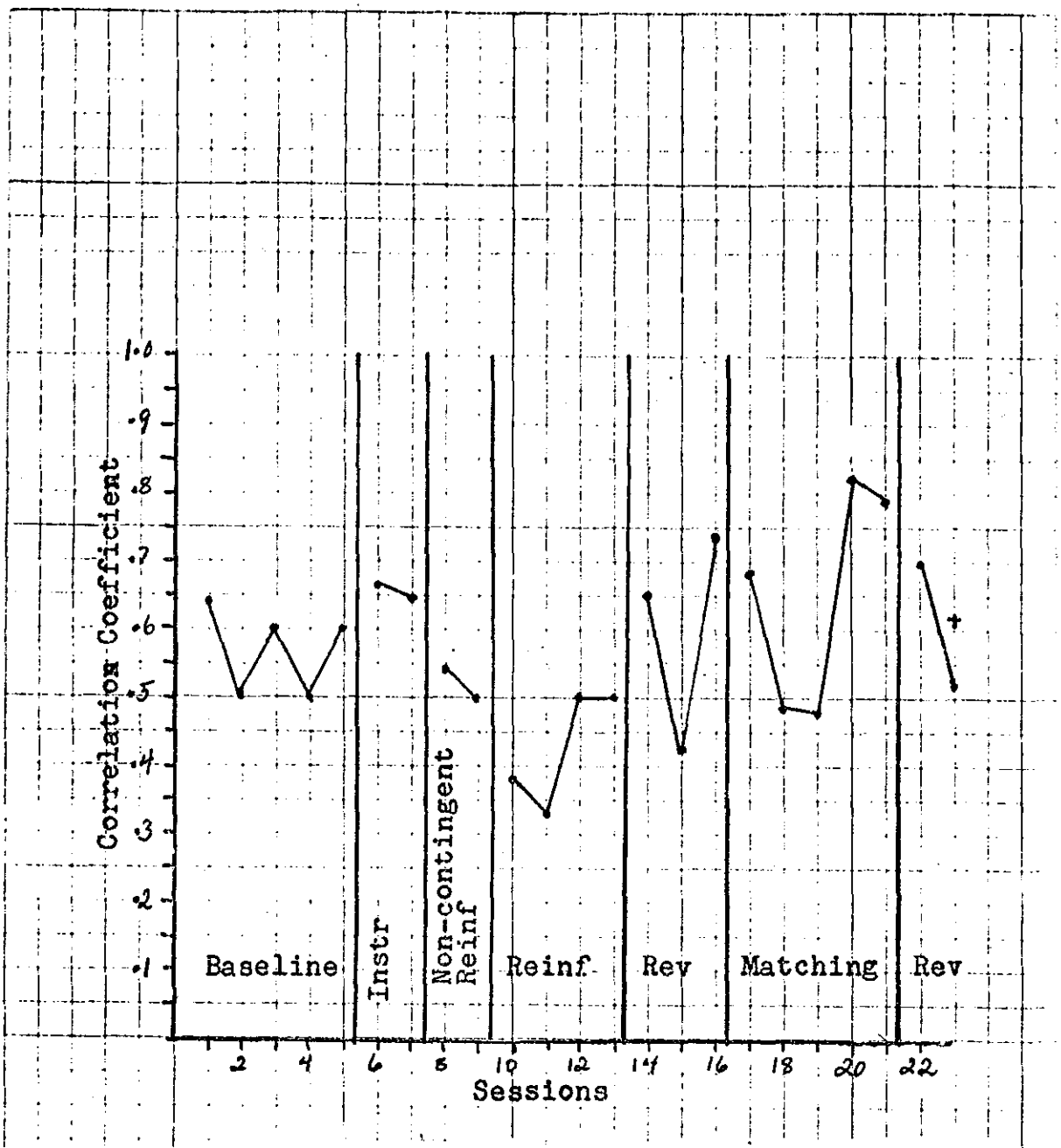


Figure 4. Linear correlation coefficients between the mean standard and the level of performance for the experimental words. (+ = 1 exp. session).

out of 15 words spelled correctly would equal what is considered satisfactory performance in most classrooms. Most students, however, set a standard of 10 correct or lower as their standard for good or excellent performance. Thus it appears that children do, in fact, need training in self-determining standards for performance.

Instructions and non-contingent reinforcement for setting higher standards, two techniques that teachers and parents often use to change behavior, were essentially ineffective in teaching children to self-determine standards of performance. Children who did change standards somewhat during these conditions were probably demonstrating individual idiosyncratic behavior rather than responding to the experimental manipulations.

Contingent reinforcement for raising standards of performance appeared to be a viable means of teaching children to set high standards for performance. The parallel increase in the level of self-determined standards for the probe words further supports the utility of this technique (Premack & Anglin, 1973). This finding seems to indicate that children might only need to be trained to set standards for one academic behavior in order to be able to set appropriate standards for performance in a similar academic task. However, the degree of generalization that can be obtained as the similarity between the tasks decreases remains to be investigated.

In evaluating the effects of contingent reinforcement on the level of self-determined standards, one must also take into account some of the individual variations that occurred in this phase. Although all

subjects raised their standards, 11 of them did not lower their standards in the subsequent reversal phase or lowered them slightly such that they still remained above baseline levels. Hence it appears that variables other than contingent reinforcement for raising standards were affecting the level of self-determined standards for these subjects.

For at least 10 of these 11 subjects, the relationship between standards and performance may provide a possible explanation for their behavior during the reversal I phase. That is, these subjects' level of performance closely matched their level of self-determined standards such that they were receiving ratings of excellent or good during the contingent reinforcement and reversal I conditions. This degree of matching may have, in and of itself, provided sufficient reinforcement for maintaining standards at this increased level.

For the subjects who maintained standards at the maximum level (14), an alternate hypothesis is also plausible. Due to the fact that the maximum level of self-determined standards was attained during the contingent reinforcement phase, the degree to which subjects may have raised standards is not apparent. In other words, subjects may have raised standards such that their mean standard exceeded 14, had this been feasible. Thus, the maintenance of standards at this maximum level during the reversal phase may in fact obscure a decrease in the level of self-determined standards, over the standards the subjects might have set had there been no ceiling. The existence of such a ceiling becomes even more probable when one considers the fact that

25 out of the 30 subjects set standards at the maximum level during the contingent reinforcement phase.

Although contingent reinforcement for raising standards seemed to be a successful procedure for training children to set high standards, the effects of high standards on the level of performance must also be considered. According to Bandura (1971) high standards should result in high levels of achievements. In other words individuals should stive to meet their standards in order to avoid the negative self-evaluations that occur when standards are not met. The results of this study do not, however, support Bandura's thesis. Two thirds of the students did demonstrate some improvements in performance during this phase, but these improvements were small relative to the increases in the level of self-determined standards. Furthermore, decreases in the level of self-determined standards in the subsequent reversal phase were not accompanied by similar decreases in performance levels. Thus, one might attribute the improvements in performance that occurred during the contingent reinforcement condition to a practice effect; that is, through continued exposure and practice, the students became more proficient in spelling.

In an attempt to develop a relationship between the level of self-determined standards and the level of performance, the matching phase was instituted. The results from this phase suggest a moderate overall increase in standards and performance for most subjects. Subjects who were already setting standards at the maximum level at the onset of this phase maintained standards at this level and demonstrated

improvements in performance by attaining a performance rating of excellent rather than good. The lowering and raising of standards by some subjects in this phase was a direct function of the matching procedure. Subjects were instructed to lower and raise their standards contingent upon their level of performance. The matching procedure seemed to be a somewhat successful technique for getting children to raise standards of performance and the level of performance simultaneously. The similar increases in standards and performance that occurred for the probe words during this phase, further supports the efficacy of this technique. The fact that these increases were maintained for many subjects when contingent reinforcement for meeting one's goal was withdrawn (reversal II) however, places a limit on the degree to which one may attribute these changes to the experimental manipulation.

The improvements in performance that occurred during the matching phase have other possible explanations apart from the matching contingency itself. Performance appeared to be improving for many subjects throughout the duration of the study and continued to improve for some in the last reversal phase. Thus, changes in the level of performance may simply reflect a practice effect. In addition, the degree to which the results obtained in the matching phase can be generalized to subjects who had not experienced the same prior experimental conditions is limited. It is possible that the changes in performance that occurred in the matching phase were a function of the preceding conditions. That is, contingent reinforcement for raising standards resulted in an increase in the level of self-determined standards. Standards were

then lowered by most subjects in the subsequent reversal phase but to a level somewhat higher than baseline. When the matching phase was introduced subjects were required to perform at a level comparable to their level of self-determined standards. Thus, this contingency coupled with the higher level of self-determined standards compared to baseline, resulted in the higher levels of performance. Should this be the case, the limits of a reversal design, which does not allow one to totally erase the effects of prior experimental conditions, become apparent.

The fact that the increases in the level of self-determined standards were maintained for many subjects in the reversal II phase also warrants explanation. It is possible that meeting one's goal or standard is sufficiently reinforcing to maintain increased levels of self-determined standards, as was suggested earlier. This would also explain why many of these subjects did lower standards after the contingent reinforcement condition. In that phase, subjects were unable to perform at a level comparable to their level of self-determined standards. However, the fact that some subjects were setting standards at the maximum level (14) during the contingent reinforcement and matching phases places an additional limitation on the interpretation of the data from this phase.

An additional point that requires some consideration is the apparent negligible effect the matching procedure had for several subjects. This occurrence can be explained by several weaknesses in the matching procedure itself. For one, subjects had to meet their goal twice in a row before they were required to raise their standards.

Some subjects met their goal on every other test and thus according to the contingencies, never had to raise their standards. Consequently improvements in performance for these subjects did not occur. Although there were bonus privileges available for those students who met their goal three times in succession, apparently these privileges were not sufficiently reinforcing for some subjects. Furthermore, some subjects set low standards relative to their level of performance at the onset of the matching phase, and could thereby meet the contingencies of this condition without necessarily improving performance.

As the results obtained during the contingent reinforcement phase of this study suggest the relative independence between changes in the level of self-determined standards and changes in the level of performance, the correlation between standards and performance for the different phases warrants some attention. In baseline almost all subjects set standards such that they always attained a rating of performance of excellent or good. Many subjects set low standards relative to their level of performance and demonstrated this type of behavior consistently in the reversal phases. Perhaps these students had a poor perception of their ability. However, based on the comments made by the students themselves ("I just want to make sure that no matter what, I get an excellent"), it seems more likely that standards for these subjects were unrelated to the amount of effort they intended to exert. That is, whether they did well on a spelling test or not, was controlled by other unknown factors in the environment rather than the standards they set for themselves.

In the contingent reinforcement phase the relationship between standards and performance changed somewhat. Due to the increases in the level of self-determined standards and the relatively unchanged levels of performance, many subjects were now setting standards above their level of performance, such that they were receiving an unsatisfactory rating of performance. The correlation coefficients for the latter part of this phase are essentially unchanged from baseline however. This was due to the fact that students who were attaining high scores on the spelling tests were now setting standards at a comparable level.

During the matching phase one would expect high correlations between standards and performance. The correlation data that was obtained for this phase has several explanations. Since the emphasis was upon matching or meeting one's goal (standards), subjects often exceeded their goal in the early part of the matching phase. In other words they had not yet determined how much effort was required to meet their goal. Apparently, the magnitude of reward was not sufficient to maintain the behavior of exceeding one's goal. Thus, in the latter part of the matching phase subjects' level of performance more closely matched their level of self-determined standards. In addition many subjects had set low standards for themselves relative to their level of performance at the onset of the matching phase. As standards increased during this condition, the gap between performance and standards diminished.

On the basis of the correlation data as a whole, it can be concluded that students will set standards at a level they can easily achieve. Contrary to Bandura's theory (1967), the children in this study did not raise their standards when they met their goal unless they

were reinforced for doing so. In other words changes in standards seems unrelated to changes in performance. Raising or lowering one's standards does not necessarily result in parallel increases and decreases in performance. However, when the contingencies are applied to both standards and performance it is possible that this type of relationship can be developed.

Although the data obtained in this study present some limitations of interpretation., several interesting findings do appear. The results seem to indicate that children need to learn how to self-determine high standards of performance, but more importantly need to learn the relationship between standards and performance. The matching condition in this study was designed to accomplish just that, but needs some refinements in order to become a more successful procedure. The matching phase was essentially a shaping procedure which is particularly difficult to execute for 30 students. For example, the privilege sheet should have been designed to satisfy the individual preferences of each child rather than organized on the basis of an average rank. The fact that the privileges held different reinforcing values for different subjects may account for some of the variability that occurred in the level of self-determined standards and the level of performance.

Also, the matching procedure could have been more effective if subjects were required to increase standards by an amount comparable to the magnitude of increase they demonstrated on the performance measure. In other words a subject who improved performance by 5 points should have raised standards by a like amount, rather than by the 1 point

minimum. In addition provisions should have been made for subjects exceeding their goal, and the bonus privileges that were available to students who met their goal three times in succession should have been more immediate. These bonus privileges did not appear to maintain the consistency in performance for which they were intended.

Along with the inherent weaknesses in the matching procedure, several external variables appeared to have minimized the effectiveness of this technique. The timing of the matching phase such that it occurred near the end of the school year presented a number of limitations. First, the matching procedure, like any other shaping procedure is a slow process and requires a great deal of time in order to be effective. This time was not available. Second, by the time the matching phase was introduced, many of the classroom privileges that had been used as reinforcers were beginning to lose their reinforcing properties. Finally, the results obtained in the second reversal phase are confounded by the same variables, in addition to representing only one to three data points for most subjects. As an evaluation of the effectiveness of the matching phase is somewhat contingent on the results obtained in this reversal phase, such an evaluation becomes even more difficult.

It appears that further research is needed to evaluate fully the degree to which such a matching procedure is an effective means of teaching children to set high standards for performance and then to meet these increased performance requirements. The degree of generalization to other untrained behavior that such a procedure appears to

produce, and the possible enduring effects that attaining one's level of self-determined standards appears to have, support the value of this technique. However, it is suggested that one not employ a reversal design in order to assess the effects of this procedure.

Appendix A

Sample Spelling List and Spelling Test Paper

Spelling List #1

* = Privilege Words

*1.	attractive	*16.	measurement
2.	disgust	17.	suffix
*3.	distant	*18.	spoon
4.	energy	19.	contagious
*5.	prepare	*20.	magazine
6.	instruction	*21.	announcement
*7.	inexpensive	22.	district
*8.	atom	23.	sour
*9.	arrangement	24.	tropic
*10.	leak	25.	public
11.	copy	26.	variety
*12.	respect	*27.	erect
13.	spoil	*28.	consist
14.	slumber	29.	identity
*15.	iris	30.	moraine

Spelling Test #1

* = privilege words

Grade for * words_____

Name_____ Grade for other words _____

*1.	16.
2.	17.
3.	18.
*4.	19.
5.	20.
6.	*21.
7.	22.
*8.	*23.
*9.	*24.
10.	*25.
11.	26.
*12.	*27.
*13.	*28.
*14.	*29.
15.	*30.

Appendix B

Sample Standard Sheets for the Different Experimental Conditions

I. Standard sheet for Baseline, Instructions, and Reversal I & II

STANDARD SHEET

Name_____

Test No.		Excellent	Good	Satisfactory	Unsatisfactory
	* WORDS				
	other words				

II. Standard sheet for Non-contingent Reinforcement and Contingent Reinforcement

STANDARD SHEET

Name_____

Points		Excellent	Good	Satisfactory	Unsatisfactory
Earned Total No.Spent No.Left	* WORDS				
Test #	other words				

III. Standard Sheet for Matching Standards and Performance

STANDARD SHEET

Name_____

Test No.		Excellent	Good	Satisfactory	Unsatisfactory
	* WORDS				
	other words				

Appendix C

List of Reinforcers

I. Reinforcers for the Reinforcement Phases

5 point privileges

1. 15 minutes free time
2. first to leave at 3:09 P.M.
3. passing back papers to the class

10 point privileges

1. correcting papers
2. Recording grades
3. free choice of seat for one period
4. 15 minutes for art work

15 point privileges

1. 10 minutes to listen to records
2. free eraser
3. free colored pencil
4. privileges for two (15 pts. each)
   a. playing checkers, cards for 10 minutes
   b. talking for 10 minutes

20 point privileges

1. privileges for two (20 points each)
   a. free choice of activity (with Mr. T's okay)
   b. talking for 15 minutes

25 point privileges

1. Certificate good for one spelling assignment
2. Free choice of seat for one day

30 point privilege

1. 25 Certificate

II. Reinforcers for the Matching Standards and Performance Phase

Excellent

1. free colored pencil
2. free choice of seat for one period
3. first to leave at 3:09 P.M.
4. first to leave for lunch
5. teaching an assignment
6. correcting papers or recording grades
7. 15 minutes for talking, playing games, art work, dictionary words, to work outside, etc.

Good

1. messenger
2. free pencil
3. first to leave at recess
4. passing papers back to the class

Satisfactory

1. 5 minutes for puzzle games
2. referee
3. board monitor
4. 10 minutes to read books, magazines, comics

Bonus Privileges

If you meet your goals for 3 spelling tests in a row you may also choose one of these privileges

1. 25 certificate
2. certificate good for one spelling assignment
3. free eraser
4. free choice of seat for the day

Appendix D

List of Privileges for the Three Ratings of Performance

Excellent Work

1. First to leave for lunch
2. 10 minutes to read books, magazines, comics
3. 15 minutes to do dictionary words
4. free pencil
5. permission to work outdoors for 15 minutes
6. messenger

Good Work

1. Teaching an assignment
2. 5 minutes free time
3. first to leave for recess
4. room monitor

Satisfactory Work

1. Board monitor
2. Paper monitor
3. 5 minutes for puzzle games
4. referee

Appendix E

Written Instructions for Raising Standards of Performance

Special Points for Raising Standards

How to earn special points for the privilege spelling words

1. If you raise your standards for one of the three ratings (excellent, good, satisfactory) you will earn one point.

Example	Excellent	Good	Satisfactory
old standards	15-13	12-9	8-5
new standards	15-13	12-10	9-5

2. If you raise your standards for two of the three ratings you will earn 2 points.

Example	Excellent	Good	Satisfactory
old standards	15-13	12-9	8-5
new standards	15-14	13-9	8-6

3. If you raise your standards for all three of the ratings you will earn 5 points.

Example	Excellent	Good	Satisfactory
old standards	15-13	12-9	8-5
new standards	15-14	13-10	9-6

4. If you have the highest possible standards and keep them that way you will earn 10 points

Example	Excellent	Good	Satisfactory
	15	14	13

5. If you lower your standards for one of the three ratings you will lose one point.

Example	Excellent	Good	Satisfactory
old standards	15-13	12=9	8-5
new standards	15-12	11-9	8-5

6. If you lower your standards for two of the three ratings you will lose 2 points.

Example	Excellent	Good	Satisfactory
old standards	15-13	12-9	8-5
new standards	15-13	12-8	7-4

7. If you lower your standards for all three of the ratings you will lose 5 points.

Example	Excellent	Good	Satisfactory
old standards	15-13	12-9	8-5
new standards	15-12	11-8	7-4

So try and raise your standards for the privilege words so you can earn lots of points and get special privileges.

Appendix F

Written Instructions for Matching Standards and Performance

New Ways to Earn Privileges

From now on, in order to be able to pick a privilege from the privilege sheet you must pick a goal. That is, you must decide whether you want to get an excellent, a good, or a satisfactory, on the next spelling test. Then you must get that rating in order to pick a privilege. For example, if you decide you want to get a good on the next spelling test you must get a good before you can pick a privilege. You may pick a rating as your goal only twice. So you could decide that you wanted to get a good on the next two spelling tests. Then you must either choose a higher rating as your goal which in this case would be an excellent, or you can raise your standards for good and choose a rating of good as your goal two more times. So if your standards for good were 9-7 you would have to raise them to 9-8. Once you pick excellent as your goal, the only thing you can do is raise your standards. So if your standards were 15-11 for excellent, you would have to raise them to 15-12. You will not be able to pick a privilege if you do not raise your goal after you have met it twice in a row, or if you lower standards or goal. So if you picked good as your goal and you got a good, you may not pick satisfactory as your goal. But if you picked good as your goal and you did not get good on two spelling tests in a row, then you can lower your goal. If you meet your goal three

times in a row you may pick a special bonus privilege too. So from now on when you set your standards you must also choose one of the ratings (excellent, good, satisfactory) as your goal for the next spelling test. You only set goals for the privilege words, not the other 15 words.

References

Baer, D. M., Wolf, M. M., & Risley, T. R. Some current dimensions of applied behavior analysis. Journal of Applied Behavior Analysis, 1968, 1, 91-97.

Bandura, A. Vicarious and self-reinforcement processes. In R. Glaser (Ed.), The nature of reinforcement. New York: Academic Press, 1971, pp. 228-278.

Bandura, A., Grusec, J. E., & Menlove, F. L. Some social determinants of self-monitoring reinforcement systems. Journal of Personality and Social Psychology, 1967, 5, 449-455.

Bandura, A., & Perloff, B. Relative efficacy of self-monitored and externally imposed reinforcement systems. Journal of Personality and Social Psychology, 1967, 7, 111-116.

Bandura, A., & Whalen, C. K. The influences of antecedent reinforcement and divergent modeling cues on patterns of self-reward. Journal of Personality and Social Psychology, 1966, 3, 373-382.

Bolstad, O. D., & Johnson, S. M. Self-regulation in the modification of disruptive behavior. Journal of Applied Behavior Analysis, 1972, 5, 443-454.

Broden, M., Hall, R. V., & Mitts, B. The effects of self-recording on the classroom behavior of two eighth grade students. Journal of Applied Behavior Analysis, 1971, 4 (3), 191-199.

Drabman, R. S., Spitalnik, R., & O'Leary, D. K. Teaching self-control to disruptive children. Journal of Abnormal Psychology, 1973, 82, 10-16.

Felixbrod, J. J., & O'Leary, D. K. Effects of reinforcement on children's academic behavior as a function of self-determined and externally imposed contingencies. Journal of Applied Behavior Analysis, 1973, 6, 241-250.

Fixsen, D. L., Phillips, E. L., & Wolf, M. M. Achievement Place: The reliability of self-reporting and peer reporting and their effects on behavior. Journal of Applied Behavior Analysis, 1972, 5, 19-30.

Glynn, E. L. Classroom applications of self-determined reinforcement. Journal of Applied Behavior Analysis, 1970, 3, 123-132.

Glynn, E. L., Thomas, J. D., & Shee, S. M. Behavioral self-control of on-task behavior in an elementary classroom. Journal of Applied Behavior Analysis, 1973, 6, 105-113.

Kanfer, F. H., & Duerfeldt, P. H. Comparison of self-reward and self-criticism as a function of types of prior external reinforcement. Journal of Personality and Social Psychology, 1968, 8, 261-268.

Kanfer, F. H., & Marston, A. R. Determinants of self-reinforcement in human learning. Journal of Experimental Psychology, 1963, 66, 245-254.

Kanfer, F. H. Behavior modification in education: A few opinions on critical issues. Submitted for publication, University of Illinois, 1973.

Locke, E. A., Cartledge, N., & Koeggel, J. Motivational effects of knowledge of results: A goal setting phenomena? Psychological Bulletin, 1968, 70, 474-485.

Lovitt, T. C., & Curtis, K. Academic response rate as a function of teacher and self-imposed contingencies. Journal of Applied Behavior Analysis, 1969, 2, 49-53.

Marston, A. R., & Kanfer, F. H. Human reinforcement: Experimenter and subject controlled. Journal of Experimental Psychology, 1963, 66, 91-94.

O'Leary, D. K., & Drabman, R. Token reinforcement programs in the classroom: A review. Psychological Bulletin, 1971, 75, 379-398.

Premack, D. & Anglin, B. On the possibilities of self-control in man and animals. Journal of Abnormal Psychology, 1973, 81, 137-151.

Santogrossi, D. A., O'Leary, D. K., Romanczyk, R. G., & Kaufman, K. F. Self-evaluation by adolescents in a psychiatric hospital school token program. Journal of Applied Behavior Analysis, 1973, 6, 277-287.