DOCUMENT RESUME

ED 039 267 UD 009 942

AUTHOR Peisach, Estelle; Victor, Jack
TITLE The Development of a Group Speech Sound Test for Disadvantaged Children. Final Report.
INSTITUTION New York Univ., N.Y. Inst. for Developmental Studies.
PUB DATE Dec 69
NOTE 56p.

EDRS PRICE EDRS Price MF-$0.25 HC Not Available from EDRS.
DESCRIPTORS *Auditory Discrimination, *Disadvantaged Youth, Elementary School Students, Group Tests, *Linguistic Performance, Lower Class, *Negro Students, Speech Education, Speech Evaluation, Speech Improvement, *Speech Tests, Test Results, Urban Youth
IDENTIFIERS Wepman Auditory Discrimination Test

ABSTRACT

Because speech sound discrimination tests can be contaminated by a particular kind of "response set," i.e., the tendency not to respond to the final parts of words as effective stimuli, children can be mistakenly diagnosed as having very poor speech-sound discrimination. About 20 percent of urban lower-class children have this response set difficulty and could benefit from an appropriate remedial program. This project has been aimed at developing a group-administered speech-sound test, appropriate for administration to young lower-class black urban children. Focus of the study has been more on identification of response set patterns rather than the evaluative criteria: subjective reaction of the testers to administration of various formats, general performance of children, ability of a particular format to discriminate between initial and final phoneme performance, and reliability measures. Detailed item analyses have been made in order to locate problem areas of discrimination. Instructions for administering the speech-sound discrimination test and illustrations thereof, and tables of test results are appended. [Not available in hard copy due to marginal legibility of the original document.] (KG)

OEO Contract No. B89-4612 (C)

Final Report, December, 1969

The Development of a Group Speech Sound Test for Disadvantaged Children

Institute for Developmental Studies
School of Education
New York University

Director: Martin Deutsch, Ph.D.

Principal Investigator: Estelle Peisach, Ph.D.

Co-Principal Investigator: Jack Victor, Ph.D.

ABSTRACT

A recent study has shown that speech-sound discrimination tests can be contaminated by a particular kind of "response set", i.e., the tendency not to respond to the final parts of words as effective stimuli. As a result of this contamination, children can be mistakenly diagnosed as having very poor speech-sound discrimination. Because speech-sound discrimination is viewed as a developmental phenomena, often remedial training is put off until the child is about eight years of age; and then, the remedial training may not focus on the child's real difficulty--the response set. It is estimated that about 20% of the population of urban disadvantaged children have this response set difficulty and could probably benefit by an appropriate remedial program.

The difficulty with developing appropriate training procedures is that the only available test that is constructed for (though not scored for) distinguishing between speech-sound discrimination difficulties and response set problems is an individually administered test. Any reasonably designed study which would deal with training procedures for children showing evidence of response sets would require the pretesting of about four hundred children. This is obviously a very costly procedure.

This project has been devoted to the development of a group-administered speech-sound test, appropriate for administration to young, lower-class, urban black children. We have been more concerned with the identification of the response set pattern than the evaluative criteria: (a) subjective reaction of the testers to administration of the various formats; (b) general performance of the children; (c) ability of a particular format to discriminate between initial and final phoneme performance; and (d) reliability measures. We also made some detailed item analyses in order to locate problem areas of discrimination in this particular population.

Through pilot-testing a variety of formats, it was evident that although speech-sound discrimination in these young disadvantaged subjects could be evaluated by group procedures, much of what is viewed as speech-sound discrimination difficulties in test protocols, may be in reality, confusion over instructions. In addition, speech-sound discrimination of initial consonant sounds is quite different from speech-sound discrimination of final consonant sounds. The relationship between the difficulty levels of the various phonemic discriminations in the two positions is either negative or zero, thus implying the need for separate measures and differential remedial procedures.

BRIEF DESCRIPTION OF THE PROJECT

Studies concerned with speech-sound discrimination (SSD) (a specific type of auditory ability thought to be highly related to speech and reading skills) generally have produced conflicting results. The inconsistencies found in these studies have been interpreted by Russell and Fea (1963) and Coller (1967) to be due, in part, to the use of measuring instruments which evaluate (more or less adequately) different, and often seemingly incompatible aspects of SSD. In addition, the results of a recent study by Coller, Coleman, and Schwartz (1967) show that the scoring procedure used for diagnosing SSD ability in one of the more popular SSD tests is contaminated by "response set" behaviors. Coller, et al., claim that the Wepman Auditory Discrimination Test (1958), a bi-positionally balanced, paired speech-sound test, is not, under its present scoring system, a valid measure of speech-sound discrimination. The suggested scoring system of the Wepman Auditory Discrimination Test (WADT) sometimes permits children who have only "response set" problems, i.e., the tendency not to treat the final parts of words as effective stimuli, to be mistakenly diagnosed as having very poor SSD ability. It should be clear that both the inaccurate diagnosis of a child's SSD ability and the use of inappropriate scoring systems can most certainly help to create inconsistencies when SSD ability is related to speech and reading skills.

It is easy to see that a lack of understanding of what speech-sound discrimination tests are actually measuring can cause gross

errors in the diagnosis of the child's speech-sound discrimination ability. What is not too fully realized is that such errors can be compounded when remediation procedures are suggested. Either inefficient, inappropriate, or no remedial programs are typically advised. For example, Wepman (1960) has suggested that, "...speech correction for children whose auditory discrimination develops slowly should not be started until after the seventh year [p. 329]." Accordingly, if a child showing evidence of response set behavior is mistakenly diagnosed as having poor SSD ability, remedial training might be put off and not started until it is too late.

It is estimated that about 20% of Wepman test protocols of the disadvantaged population are mistakenly diagnosed as indicating that the children have poor speech-sound discrimination ability. Coller, et al., claim that the majority of such children really have a response set problem, that is, attending to initial sounds, but not attending to the same sound when located in a final position. This indicates a need for a test which can be administered to groups of children so that an accurate estimate of the child's SSD ability can be quickly and easily obtained. Such a test should also be able to determine if the child has a response set problem.

This project has been devoted to the development of a group-administered speech-sound test, appropriate for administration to young, lower-class, urban black children. We have been more concerned with the identification of the response set pattern than the evaluation of overall speech-sound discrimination although this test could provide a measure of this ability. We, therefore, developed a test which could give us separate measures for the discrimination

of the same phonemes in initial and final positions of minimal phonemic difference pairs.

The year's work involved the selection and preparation of items and the investigation of appropriate administration procedures. We were particularly concerned in regard to two problems: (1) that neither the format of the test nor the instructions should be a barrier to obtaining a reliable estimate of the child's ability to discriminate particular phonemes; (2) that the content of the items not be unduly unfamiliar for children of this age and background, again, so that the measure attained should reflect speech-sound discrimination ability rather than some other variable such as the extent of the child's vocabulary.

We consequently pilot-tested a variety of formats using several evaluative criteria: (a) subjective reaction of the testers to administration of the various formats; (b) general performance of the children; (c) ability of a particular format to discriminate between initial and final phoneme performance; and (d) reliability measures. We also made some detailed item analyses in order to locate problem areas of discrimination in this particular population.

METHODOLOGY

Work on this project proceeded through several stages. The selection of items and their illustration; preparation of variations of the same test, differing only in format or administration procedures; administration of the various formats to pilot samples; examination of the results with subsequent selection of three formats for further investigation; administration of the selected forms to new samples of similar subjects and analysis of the data obtained on the three selected formats.

Because of our great concern to find formats and administration procedures appropriate for use with young, urban, disadvantaged children, we developed five versions of the group administered speech-sound test. In addition, for the same reason, great care was exercised in the selection of items and their illustrations.

Each test item required that a minimal phonemic discrimination be made. For Formats 1 and 2, discrimination had to be made between illustrations of two words, e.g., a picture of a gate and a picture of a date. In Format 1, the child had to mark the picture that contained the sound voiced by the tester whereas in Format 2, the child had to mark the picture of the word said by the examiner. For the three other formats (3, 4 and 5), discrimination had to be made between pairs of stimuli, e.g., gate/gate and gate/date. Of these three formats, the first two contained literal representations of the words, whereas Format 5 merely presented two symbols which were the same (two circles) versus two symbols which differed

(a circle and a square). In Format 3, the child was told to mark the words said by the examiner. In Format 4, there was an additional instruction advising the child to mark the pair that was the same or different according to whether the words said by the examiner were the same or different. These last instructions were also used with the pictures of symbols. These latter three formats were developed in the hope that the additional clue regarding the sameness or difference would counteract any lack of familiarity with the stimuli. If use of the symbols were understood by the children, this would have vastly extended the possible range of test items.

Since the response set phenomenon (correct response to a particular phoneme in the initial position but incorrect response to the same phoneme in the final position) was obtained with the individually administered Wepman Auditory Discrimination Test, the phonemes used in each word-pair were derived from the Wepman List. The Wepman List includes 13 paired phonemes: (1) three voiced stops (g/b, d/b, g/d); (2) three unvoiced stops (k/p, t/p, k/t); (3) one voiced fricative (~~th~~); (4) five unvoiced fricatives (f/p, f/s, s/~~th~~, s/sh, ~~th~~/sh); and (5) one nasal (m/n). Our test contains eleven paired phonemes. Three paired phonemes were eliminated because it was impossible to find words to fit them which could be pictorially represented and recognizable by children. An r/l paired phoneme was added to our list, as this pair seemed to be one of common difficulty in discrimination for this population.

Each of the eleven phoneme-pairs appeared both at the beginning of a word-pair and at the end of a word-pair (e.g., cot/pot and cake/cape). Altogether there are twenty-two different word-

pairs. For these twenty-two items, in the single item discrimination tests, the correct phoneme was found only in one of the two presented items. In the paired item discrimination tests, the correct response was the pair of different stimuli. In addition, for the latter tests there were also eleven items in which the correct response was the word-pair that was the "same". These additional eleven items were randomly chosen from the words which were included in the "different" items. For the additional eleven single discrimination items, the phoneme not selected in the original item was the correct choice. Position of the correct response was randomly assigned throughout each test. The items are presented in Tables 6, 7 and 8 in the Findings Section where they are discussed with reference to the difficulty level of the various phonemic discriminations.

A total of 84 Black and Puerto Rican subjects judged to be lower SES on the basis of attendance in a selected school received either two or three of the five varying test formats. The tests were administered in reading classes which met for forty minutes a day. Once time for the children to arrive and settle down and for the testers to set up was taken into account, it was found that only one test could be administered comfortably in one period. Therefore, each child was tested in two or three sessions. As a result, control over order of presentation, and assignment of subjects of equal ability to the various test formats was not feasible. Further, because of absences, different numbers of children were administered the various test formats.

Because of the assignment of reading groups organized by the school for testing, the size of the groups varied considerably.

The examiners found that because of the organization of the test booklet, it was fairly easy for the children to follow and complete the test. The largest group tested consisted of 13 subjects. There was no difficulty with a group of this size. The examiners estimated that administration of this test by two examiners tc ten to fifteen children (half a class) is very comfortable. We, therefore, eliminated the phase of this study which was to be concerned with determining the relative size of the group to which such a test could be successfully administered.

Each child was presented a booklet containing thirty-five pages with one test item on each page. There were two sample items which were drawn on the blackboard as well as appearing in the booklet, and which were administered directly by the tester. The purpose of the sample items was to orient the child to the presentation format of the booklet and to make sure he understood the instructions. To avoid distraction, each test item appeared on a separate page. To aid the subject in keeping his place in the test booklet, a picture such as a ball or a flower (in the manner of the Lorge Thorndike Intelligence Test--Level 1) was placed on the upper right-hand corner of each page. The children were instructed to raise hands if the proper picture did not appear on the page, so that a tester could be of assistance. During testing, there were always at least two testers present, one to read the items and one to circulate around the room to aid those subjects who were having difficulty for one reason or another.

In order to evaluate the various formats as to their appropriateness for the purpose for which they had been designed, we considered both the testers' subjective reactions to administering

the various formats as well as more formal statistical characteristics of group results.

Development of a test of this type requires a somewhat unorthodox approach to evaluation of test characteristics. Essentially, the test is to be used for diagnostic purposes, that is, to detect those subjects who have a particular response set in auditory discrimination. Such an instrument, which measures perceptual discrimination, is not subject to the same degree of developmental influence that an instrument measuring a learning or cognitive skill would be. Phonemes are discrete rather than continuous qualities. One is limited, therefore, in trying to manipulate the stimuli to increase similarity between the different sounds in order to make discrimination more difficult.

Therefore, the test should not be so designed as to obtain maximum variance among scores being administered to a group of subjects. Rather, most children should be able to obtain a high if not perfect score on this test. Further, because of the decreased variance in such a test, internal consistency measures of reliability may be misleading.

In the actual evaluation of the test formats, we generally took the following position. High error scores were interpreted as representing poor administration instructions--not inability of the subjects to discriminate and not improved test discrimination. Because of the specific diagnostic purpose of this test, we looked most favorably at those formats which discriminated well between initial and final phoneme performance. At this point in our evaluation, the internal consistency measures of reliability had little

influence on our approval or disapproval of a particular format.

Format 1 was discarded because of both quantitative and qualitative objections. The testers felt that the children found the task confusing. In addition, some teachers commented that this was an inappropriate testing procedure since the instructions required behavior contrary to approved teaching procedures for reading. The fact that this test had the second highest mean error scores, 3.58 for initial phonemes and 5.15 for the final phonemes (Table 1), seems to confirm the test administrators' judgment of the inadequacy of this format.

Format 2 was most liked by the test administrators. Although the Kuder-Richardson reliabilities coefficients (Table 2) were not satisfactory, the mean error scores were low and there was good discrimination between initial and final phoneme performance. We, therefore, decided to obtain additional data on this format.

Format 3 was also chosen by the examiners as being very appropriate for the sample of children tested. In the case of this variation, the Kuder-Richardson reliabilities (Table 2) are very good, but the mean error scores are relatively high and discrimination between initial and final phonemes performance is relatively minor (Table 1). In the case of this format also, we decided to gather additional data.

Format 4 was disliked by the testers. The addition of the same-different instructions seemed to be redundant and confusing and probably not attended to by the subjects. However, despite the low reliability coefficients, the mean error scores were the lowest and the discrimination between initial and final phoneme performance

was as good or better than the discrimination obtained with any of the other formats (Table 1). We, therefore, decided to include this variation for further investigation despite the administrators' strong dislike of it.

Format 5 was disliked by the examiners. The testers felt that the children were unable to understand the instructions. The high mean error scores and poor discrimination between initial and final phoneme performance seemed to bear out the examiners' judgment (Table 1).

In summary, we discarded Formats 1 and 5, and retained Formats 2, 3 and 4 for further examination.

At the conclusion of the initial pilot-testing, the investigators were not certain that the order of presentation of tests to the same children and the relative ability of children assigned to given formats had not confounded the results. To control for these factors, each of the three formats selected for further investigation was administered to an independent sample of low SES, black, first-grade subjects. Seventy subjects were included in each sample. Each sample was tested twice with one week interval between testing in order to obtain test-retest reliability data. Because of absences and invalid protocols, the size of the final samples were 57 for Format 2, 52 for Format 3, and 44 for Format 4. This time the formats were put on tape and played to each group of subjects. This procedure insured more uniform administration of the test. However, as in the first pilot-testing, the sample items were administered by the examiner. The subjects were assigned to the different formats by randomly dividing each first-grade class in half. Each

half of any class (ten to fifteen children) received a different format. The format selection was determined beforehand in such a manner that each combination of two formats occurred with equal frequency. By dividing the classes in this way, chances of obtaining essentially equal samples across the three treatments with respect to speech-sound discrimination ability were improved.

The data obtained from this second sample was then more thoroughly examined. In addition to the calculation of means and standard deviations for the initial and final phoneme subtests of the three formats (Table 3), test-retest (Table 4), and Kuder-Richardson reliability coefficients (Table 5) were calculated. Item analysis to determine the difficulty level of the items over the three formats was completed (Tables 6, 7 and 8). Consistency of item difficulty from initial to final position within each format as well as consistency of item difficulty across formats was estimated by rank order correlation coefficients (Tables 9, 10 and 11).

FINDINGS

Reported in this section are the results of the second pilot testing with the three selected formats (2, 3 and 4).

From Table 3, containing Ns, means and standard deviations for the initial and final phoneme subtests for all three formats, it can be seen that maximum discrimination between initial and final phoneme performance was obtained with Format 3, both for the initial and repeated administrations of the test. Format 3 had the highest reliabilities coefficient associated with it--both the preferred test-retest reliability coefficient, r =.76 (Table 4); and the Kuder-Richardson internal consistency reliability measure, r =.69 (Table 5). These findings, combined with the examiners' subjective evaluations, have led us to select Format 3 as appropriate for administration to young, disadvantaged children.

However, the results are somewhat less than ideally satisfactory; the generally high means indicate that the subjects were hitting the ceiling of the test. The low reliability coefficients reflect the lack of variance and ceiling effect. As discussed in the Methodology Section, the cause of this may be, to some extent, inherent in the nature of the test. Making a phonemic discrimination is an all-or-none matter, probably achieved at relatively young ages. Nevertheless, the data from all three formats was examined in greater detail to gain more understanding of what had occurred.

Essentially, the final phoneme subtest carried the entire test. From Tables 4 and 5, it can be seen that the reliability coefficients

obtained with the final phoneme subtest are almost equal in value to those obtained with the entire test.

Tables 6, 7 and 8 present the phonemic discriminations required of the subjects and the percentage of subjects making correct responses for each discrimination as it appeared in the initial and final positions for the three formats (2, 3 and 4, respectively).

In Table 9, the percentage correct for each discrimination in both positions has been averaged across test administrations and repetitions of the discriminations, and then ranked.

From the correlations reported in Table 10, it can be seen that the rank positions of both the initial and final phonemic discriminations remain fairly constant from one test to another.

Examination of Table 11, coefficients of the rank difficulty between phonemic discriminations in initial and final positions across test-retest administrations of the three formats, reveals consistent substantial positive correlations for both the initial and final phoneme subtests. However, within any given administration the correlation between the initial and final phoneme subtests is either negative or negligible. This explains the lack of improvement in total test reliability which one would expect from the increased number of items involved. We seem to have two quite different tests, one considerably more difficult than the other.

Although aware that generally none of the discriminations were very difficult for the children to make, we, nevertheless, examined Tables 6 through 9 to determine which phonemic sounds were responsible for most of the variance of the test. Although their ranks

within the first five positions across the formats do change, consistently, the five most difficult initial discriminations are: g/d, ch/sh, m/n, sh/s and s/f. One pair, p/c, was found difficult in Format 2, but easy in Formats 3 and 4. In final position, the five most difficult sounds are: p/c, d/b, g/d, sh/s and g/b. Two inconsistencies among formats occurred: t/p was second in difficulty in Format 3; and r/l was second in Format 4. No explanation is found for these inconsistencies. It appears that g/d and sh/s are among the most difficult in both initial and final positions.

Since there is only one item per phonemic discrimination in each position, it is difficult to determine from the tests whether or not some difficulty may be related to the phonemic context, that is, the possible differences in vowel sounds across varying stimuli, or even, the general familiarity of the items. Specific reactions to particular stimuli may also account for variations in difficulty level. The examiners believe that the difficulty of the s/f discrimination may have been due to the children's amused reaction to the illustration of "feet," with consequent lack of attention to what the examiner (or tape) was saying. The examiners further believe that the difficulty associated with the g/d discrimination in the initial position may be related to its being the first item of the test. The rhyming nature of the test items, which is not so for the sample items, may at first confuse the subjects.

CONCLUSIONS

It is possible to develop a group-administered speech-sound discrimination test appropriate for administration to young disadvantaged children. (Such a test should be administered by two testers to groups ranging up to approximately 15 in number.) This test can differentiate between the child's ability to distinguish the same sounds in initial and final positions, which is important information for remedial procedures.

From our work with various formats, it would appear that much of what is viewed as speech-sound discrimination difficulties in test protocols may be, in reality, confusion over instructions. When the instructions become clearer, the group performs at a higher level.

Judging by data from previous work at the Institute with similar samples, the test described above, despite being group-administered, is easier than the Wepman Auditory Discrimination Test, an individually administered instrument. One possible explanation is that, by discarding the same-different instructions, a confounding factor to the measurement of children's speech-sound discrimination ability may have been eliminated. Evidence supporting this contention can be seen in the results with Format 4, which was identical to Format 3 except for the inclusion in Format 4 of same-different instructions. Children taking Format 3 obtained higher scores than those taking Format 4.

From our data it seems that, at least for the samples we used, speech-sound discrimination in initial consonant sounds is quite

different from speech-sound discrimination in final consonant sounds. Different phonemic discriminations are difficult as a function of position in the word. This has great implication for remedial techniques, as well as test construction of speech-sound discrimination tests. Different phonemes should be focussed upon, depending upon the context in which they appear. To achieve highly reliable measures of this ability, it is probable that separate tests or subtests should be administered for discrimination within each position. One cannot assume that adding another item with a phonemic discrimination in another position is equivalent to increasing the number of items of similar nature.

The test developed in this project reflects this problem. To improve the general reliability of this test it would be essential to add more items with the same phonemic discriminations to each subtest. Although this test was patterned after the Wepman, other sources of information regarding phonemic discrimination problem areas for this population (e.g., descriptive linguistic studies) could be drawn on as guides for additional items which would increase the diagnostic utility of the test.

TABLE 1

Ns, Mean Error, Initial and Final Phoneme Scores and Standard Deviations for Five Variations cf a Group Speech Sound Discrimination Test

		Initial		Final	
Format	N	$\bar{X}$	S.D.	$\bar{X}$	S.D.
1	48	3.58	2.98	5.15	3.23
2	29	1.10	1.26	2.83	1.90
3	41	2.37	3.22	2.80	2.93
4	24	.75	.94	1.88	1.35
5	35	4.94	3.64	5.29	3.07

TABLE 2

Kuder-Richardson #20 Reliability Coefficients for Initial and Final Phoneme Subtests of Three Variations of a Group Speech Sound Discrimination Test

Format	N	Initial	Final
2	29	.65	.69
3	41	.93	.88
4	24	.53	.55

TABLE 3

Ns, Means and Standard Deviations for Initial and Final Phoneme Subtest Scores and Total Test Scores for Three Formats of a Group Administrated Speech Sound Test

		Initial		Final		Total Test	
Format	N	$\bar{X}$	S.D.	$\bar{X}$	S.D.	$\bar{X}$	S.D.
				Test			
2	57	15.89	1.02	13.28	1.86	29.18	2.40
3	52	15.38	1.42	11.75	2.61	27.10	3.23
4	44	13.95	2.22	10.98	2.51	24.93	5.09
				Retest			
2	57	15.56	1.28	13.47	1.96	29.04	2.82
3	52	15.60	1.60	12.94	2.28	28.54	3.27
4	44	14.86	1.50	13.79	1.91	28.34	2.90

Maximum Possible Scores

Initial Phonemes	(17)
Final Phonemes	(16)
Total Test	(33)

TABLE 4

Pearson Product Moment Test-Retest Reliability Coefficients for Three Formats of a Group Administered Speech Sound Test

Format	N	Initial Phonemes	Final Phonemes	Total Test
2	57	.20	.50	.50
3	52	.21	.69	.76
4	44	.10	.44	.40

TABLE 5

Kuder-Richardson #20 Reliability Coefficients for Two Administrations of Three Formats of a Group Administered Speech Sound Test

		First Administration:			Second Administration:		
Format	N	Initial	Final	Total Test	Initial	Final	Total Test
2	57	.13	.51	.56	.22	.57	.64
3	52	.32	.61	.61	.53	.59	.69
4	44	.59	.54	.83	.29	.53	.61

TABLE 6

Percentage of Subjects Obtaining Correct Phoneme Discrimination in Initial and Final Positions

Format 2:

Sound Discrimination	Initial Position	Administration 1st	2nd	Av.	Final Position	Administration 1st	2nd	Av.
g/d	gate/date	58%	70%	64%	bud/bug	68%	88%	78%
t/c	cop/top	96%	93%	94.5%	bat/back	96%	89%	92.5%
	cop/top	96%	96%	96%	bat/back	95%	95%	95%
ch/sh	chop/shop	82%	82%	82%	wash/watch	98%	98%	98%
g/b	boat/goat	100%	96%	98%	tug/tub	74%	81%	77.5%
	goat/boat	100%	98%	99%	tub/tug	82%	82%	82%
p/c	pot/cot	95%	93%	94%	cape/cake	91%	84%	87.5%
	pot/cot	98%	84%	91%	cake/cape	19%	23%	21%
n/m	mail/nail	91%	91%	91%	gun/gum	98%	91%	94.5%
					gun/gum	91%	89%	90%
sh/s	shack/sack	95%	91%	93%	gas/gash	75%	82%	78.5%
t/p	tool/pool	100%	93%	96.5%	cap/cat	96%	95%	95.5%
	tool/pool	96%	95%	95.5%				
d/b	beer/deer	95%	91%	93%	robe/road	77%	75%	76%
	deer/beer	98%	95%	96.5%				
s/f	seat/feet	96%	98%	97%	laugh/lass	86%	90%	88%
					laugh/lass	89%	96%	92.5%
r/l	lock/rock	95%	98%	96.5%	fire/file	96%	95%	95.5%
	rock/lock	96%	89%	92.5%				

TABLE 7

Percentage of Subjects Obtaining Correct Phoneme Discrimination in Initial and Final Positions

Format 3:

Sound Disc.	Initial Position	Administration 1st	2nd	Av.	Final Position	Administration 1st	2nd	Av.
g/d	gate/gate gate/date	71%	98%	84.5%	bug/bug bud/bug	67%	52%	59.5%
t/c	top/cop cop/cop	94%	79%	86.5%	bat/back bat/bat	79%	98%	88.5%
	top/top cop/top	98%	98%	98%	back/back back/bat	92%	85%	88.5%
ch/sh	chop/chop chop/shop	77%	90%	83.5%	watch/watch wash/watch	86%	85%	85.5%
g/b	goat/goat goat/boat	94%	100%	97%	tug/tub tub/tub	48%	75%	61.5%
	boat/goat boat/boat	100%	88%	94%	tug/tug tub/tug	77%	85%	81%
c/p	cot/cot cot/pot	98%	94%	96%	cake/cape cake/cake	33%	52%	42.5%
	pot/cot pot/pot	90%	94%	92%	cape/cape cape/cake	58%	63%	60.5%
m/n	mail/nail nail/nail	85%	81%	83%	gum/gun gun/gun	98%	90%	94%
					gum/gum gun/gum	98%	96%	97%
s/sh	sack/sack shack/sack	79%	85%	82%	gas/gas gas/gash	54%	90%	72%
t/p	tool/tool tool/pool	94%	96%	95%	cap/cat cat/cat	54%	62%	58%
	tool/pool pool/pool	94%	98%	96%				
d/b	deer/deer deer/beer	98%	98%	98%	robe/road robe/robe	79%	77%	78%
	beer/deer beer/beer	92%	88%	90%				
s/f	seat/feet feet/feet	75%	73%	74%	lass/laugh laugh/laugh	86%	93%	89.5%
					lass/lass lass/laugh	88%	92%	90%
l/r	lock/lock lock/rock	100%	98%	99%	fire/fire fire/file	77%	90%	83.5%
	rock/lock rock/rock	98%	94%	96%				

TABLE 8

Percentage of Subjects Obtaining Correct Phoneme Discrimination in Initial and Final Positions

Format 4:

Sound Disc.	Initial Position	Administration 1st	2nd	Av.	Final Position	Administration 1st	2nd	Av.
g/d	gate/gate gate/date	45%	73%	59%	bug/bug bud/bug	77%	59%	68%
t/c	top/cop cop/cop	82%	91%	86.5%	bat/back bat/bat	70%	100%	85%
	top/top cop/top	77%	93%	85%	back/back back/bat	79%	80%	79.5%
ch/sh	chop/chop chop/shop	66%	93%	54.5%	watch/watch wash/watch	73%	100%	91.5%
g/b	goat/goat goat/boat	100%	100%	100%	tug/tub tub/tub	50%	73%	61.5%
	boat/goat boat/boat	100%	93%	90.5%	tug/tug tub/tug	86%	80%	83%
c/p	cot/cot cot/pot	95%	91%	93%	cake/cape cake/cake	39%	59%	49%
	pot/cot pot/pot	100%	91%	95.5%	cape/cape cape/cake	43%	45%	44%
m/n	mail/nail nail/nail	64%	59%	61.5%	gum/gun gun/gun	95%	98%	96.5%
					gum/gum gun/gum	95%	100%	97.5%
s/sh	sack/sack shack/sack	69%	80%	74%	gas/gas gas/gash	39%	77%	55%
t/p	tool/tool tool/pool	95%	93%	94%	cap/cat cat/cat	68%	93%	80.5%
	tool/pool pool/pool	89%	91%	90%				
d/b	deer/deer deer/beer	82%	100%	91%	robe/road robe/robe	59%	86%	72.5%
	beer/deer beer/beer	93%	93%	93%				
s/f	seat/feet feet/feet	61%	55%	58%	lass/laugh laugh/laugh	79%	95%	87%
					lass/lass lass/laugh	68%	89%	78.5%
l/r	lock/lock lock/rock	84%	95%	89.5%	fire/fire fire/file	68%	93%	50.5%
	rock/lock rock/rock	93%	95%	94%				

TABLE 9

Rank Order Difficulty Level of Phonemic Discriminations (Averaged Across Test and Retest Administrations) in Initial and Final Positions for Three Formats

Phonemic Discrimination	Rank Difficulty Order: Initial Position Formats			Final Position Formats		
	2	3	4	2	3	4
g/d	1	5	3	3	3	4
t/c	8	6	6	8	9	8
ch/sh	2	4	1	11	8	10
g/b	11	9.5	11	5	4	5
p/c	4	7.5	10	1	1	1
n/m	3	3	4	7	11	11
sh/s	5	2	5	4	5	3
t/p	9	9.5	8.5	9.5	2	7
d/b	7	7.5	8.5	2	6	6
s/f	10	1	2	6	10	9
r/l	6	11	7	9.5	7	2

TABLE 10

Spearman-Brown Rank Order Correlation Coefficient, Between Difficulty Levels of Phonemic Discriminations in Initial and Final Positions Among Three Formats

Format	Formats 3		4	
	Initial	Final	Initial	Final
2	.30	.44	.48	.52
3	--	--	.77	.71

TABLE 11

Spearman-Brown Rank Order Correlation Coefficients Between Difficulty Levels of Initial and Final Phonemic Discriminations Across Test-Retest Administration of Three Formats

Format	Rank Order Correlation Coefficients	
	I 1st Adm x I 2nd Adm	F 1st Adm x F 2nd Adm
2	.67	.89
3	.24	.71
4	.69	.39
	I 1st Adm x F 1st Adm	I 2nd Adm x F 2nd Adm
2	-.06	.13
3	-.20	-.83
4	-.63	.01
	I_{AV} x F_{AV}	
2	.08	
3	-.51	
4	-.54	

References

Anderson, P. The relationship of normal and defective articulation of the consonant (S) in various phonetic contexts to auditory discrimination between normal and defective (S) productions among children from kindergarten through fourth grade. State University Iowa. 1949. (Master's Thesis)

Aungst, L. F. & Frick, V. Auditory discrimination ability and consistency of articulation of /r/. Journal of Speech and Hearing Disorders. 1964, 29, 76-85.

Ayer, G. W. Auditory discrimination test based on Spanish. Modern Language Journal, 1960, 44, 227-230.

Biggy, M. V. Relative order of difficulty of word elements in auditory discrimination. Boston University. 1946. (Master's Thesis)

Black, J. W. & Haagen, C. H. Multiple-choice intelligibility tests, Forms A and B. Journal of Speech and Hearing Disorders, 1963, 28, 77-86.

Christine, D., & Christine, C. The relationship of auditory discrimination to articulation defects and reading retardation. Elementary School Journal, 1964, 65, 97-100.

Clark, A. D., & Richards, Charlotte J. Auditory discrimination among economically disadvantaged and nondisadvantaged preschool children. Exceptional children, 1966, 33, 259-262.

Cohen, J. H. & Diehl, C. F. Relation of speech-sound discrimination ability to articulation-type speech defects. Journal of Speech and Hearing Disorders, 1963, 28, 187-190.

Coller, A. Summary of Verbal Survey Wepman Auditory Discrimination Test: Analysis of Variance. Institute for Developmental Studies, Research Report, 1965 (a).

Coller, A. Exploratory Reliability Studies Within and Between Parallel Forms of the Wepman Auditory Discrimination Test: A First Grade Enrichment and Control Sample. Institute for Developmental Studies, Research Report, 1965 (b).

Coller, A. A Descriptive Review of Some Speech-Sound Based Tests of Auditory Discrimination. Institute for Developmental Studies, draft, 1967.

Coller, A., Schwartz, Sol, & Coleman, R. Auditory Discrimination Tests, I: Analysis of Initial and Final Phoneme Changes in Word-Pairs. Institute for Developmental Studies, Research Report, 1965.

Deutsch, C. Auditory discrimination and learning: social factors. Merrill-Palmer Quarterly, 1964, 4, 277-296.

Deutsch, M. The disadvantaged child and the learning process. In H. Passow (Ed), Education in depressed areas. New York: Teachers College Bureau of Publications, 1963.

Durrel, D. D. & Murphy, H. A. The auditory discrimination factor in reading readiness and reading disability. Education, 1953, LXXIII, 556-560.

Farquhar, M. S. Prognostic value of imitative and auditory discrimination tests. Journal of Speech and Hearing Disorders, 1961, 26, 342-347.

Gates, A. I. The Improvement of Reading. New York: Macmillan, 1927. (not read)

Hall, M. Auditory factors in functional articulatory speech defects. Journal of Experimental Education, 1938, 7, 110-132.

Hansen, B. F. The application of sound discrimination tests to functional articulatory defectives with normal hearing. Journal of Speech Disorders, 1944, 9, 347-355.

Harrington, M. J., Sister & Durrell, D. D. Mental maturity versus perception abilities in primary reading. Journal of Educational Psychology, 1955, 46, 375-380.

Hirsch, I. J., Davis, H., Silverman, S. R., Reynolds, E. G., Eidert, E., & Benson, R. W. Development of materials for speech audiometry. Journal of Speech and Hearing Disorders, 1952, 17, 321-337.

Hutton, C., Curry, E. & Armstrong, M. B. Semi-diagnostic test materials for aural rehabilitation. Journal of Speech and Hearing Disorders, 1959, 319-329.

Katz, P. A., & Deutsch, M. Visual and auditory efficiency and its relationship to reading in children. Cooperative Research Project., No. 1099, 1963.

Kronvall, E. L. An investigation of some of the factors frequently suggested as causes of functional articulation disorders. Dissertation Abstract, 1966, 26(8), 4810.

Kronvall, E. L., & Diehl, C. F. The relationship of auditory discrimination to articulatory defects of children with no known organic impairment. Journal of Speech and Hearing Disorders, 1954 19, 335-338.

Labov, W., Cohen, P., & Robins, C. A preliminary study of the structure of English used by Negro and Puerto Rican speakers in New York City. Cooperative Research Project No. 3091, U. S. Department of Health, Education and Welfare.

Lichtenberg, F. S. A comparison of children's ability to make speech sound discriminations. The Volta Review, 1966, 68, 426-434.

Mase, J. Etiology of articulatory speech defects; a comparison of the incidence of six selected factors in children having articulatory speech def cts. No. 921, Teachers College Contributions to Education, New York: Columbia University, 1946.

Murphy, H. A. Group Test for Auditory Discrimination (unpublished test) Boston University, 1941.

Murphy, H. A. Evaluation of Specific Training in Auditory and Visual Discrimination on Beginning Reading, Boston University, (Doctor's dissertation, 1943). (not read) (From Biggy, M. Virginia. Relative Order of Difficulty of Word Elements in Auditory Discrimination. Boston University (Master's Thesis, 1946).

Murphy, H. A., & Durrell, D. D. Reading Readiness Analysis (administration manual) New York: Harcourt, Brace & World, Inc., 1965.

Prins, D. Relations among specific articulatory deviations and responses to a clinical measure of sound discrimination ability. Journal of Speech and Hearing Disorders, 1963, 28, 382-388.

Pronovost, W., & Dumbleton, C. A picture-type speech sound discrimination test. Journal of Speech and Hearing Disorders, 1953, 18 258-266.

Reid, G. The etiology and nature of functional articulatory defects in elementary school children. Journal of Speech Disorders, 1947, 12, 143-150.

Reynolds, M. C. A study of the relationships between auditory characteristics and specific silent reading abilities. Journal of Education Research, 1953, 46, 430-449.

Russell, D. H., & Fea, H. R. Research on teaching reading. In Gage, N. L. (Ed.), Handbook of Research on Teaching. Chicago: Rand McNally and Company, 1963.

Scheifelbusch, R. L., & Lindsey, M. J. New test of sound discrimination. Journal of Speech and Hearing Disorders, 1958, 23, 153-159.

Schwartz, S., & Coller, A. Preliminary Analysis of Wepman Auditory Discrimination Test Data from the Study Patterns of Perceptual, Language, and Intellective Performance in Children With Cognitive, Deficits. Institute for Developmental Studies, Research Report, 1985.

Schwartz, S., Coller, A., & Coleman, R. Phonemic Discrimination and Reading Ability: An Analysis of Data From Two Institute Studies. Institute for Developmental Studies, Research Report, 1965.

Templin, M. A study of sound discrimination ability of elementary school pupils. Journal of Speech Disorders, 1943, 8, 127-132.

Thompson, B. B. A longitudinal study of auditory discrimination. The Journal of Educational Research. 1963, 56, 7, 376-378.

Trevis, L. E., & Rasmus, B. The speech sound discrimination ability of cases with functional disorders of articulation. Quarterly Journal of Speech, 1931, 17, 217-226.

Webster, D. H, & Canonge, E., eds. A teacher's guide for teaching English to native children of Alaska. (Eskimo and Athapaskan). College, Alaska: Alaska Rural School Project, U. of Alaska, 1968.

Wepman, J. Auditory Discrimination Test: Manual of Directions. Chicago: Language Research Associates, 1958.

Wepman, J. Auditory discrimination, speech, and reading. Elem. School J., 1960, 60, 325-334.

Wheeler, R. R., & Wheeler, D. A study of the relationship of auditory discrimination to silent reading abilities. J. Educ. Res., 1954, 58, 103-113.

Young, R. English as a second language for Navajos: An overview of certain cultural and linguistic factors. Window Rock, Arizona: Navajo Area Office, 1967.

APPENDIX

SPEECH-SOUND DISCRIMINATION TEST

Administration Diagram

FORMAT 3

P = Procedure L = Language

(P) First, seat the children far enough apart so that they are not able to copy easily from each other. We have tested up to 15 at a time, but you may do more if you have space and they can all hear the tape recorder. Make sure they all have a pencil or crayon. Hand out the booklets and have them each write their first and last name on the front, helping those who can't write their names.

(L) "WE ARE GOING TO PLAY A GAME WITH PICTURES. FIRST, WE WILL HAND OUT THESE BOOKLETS, AND THEN WE WILL TELL YOU HOW TO PLAY."

Testing:

1. (P) Draw two boxes on the blackboard, the first containing a star and a box, the second, a box and a box. (Same as page i in booklet.)

2. (L) "OPEN YOUR BOOKS TO THE FIRST PAGE. IT SHOULD LOOK LIKE THIS." (pointing to the board)

3. (L) "FIRST IS A PICTURE OF A STAR AND A BOX." (point to first picture), "NEXT IS A PICTURE OF A BOX AND A BOX." (point to second picture) "PUT AN 'X' ON THE ONE I SAY. BOX, BOX."

(P) Now, pick a child to come up and show you which one is correct. Then ask if everyone put an X on that one. They will be confused about what kind of X to make, so at this point, tell them they should "MAKE A BIG X LIKE THIS": and you make a big X covering the entire box containing box, box.

4. (P) Erase this example and this time drawn two boxes, the first containing a box and a star, the second, a star and a star. (Same as page ii in booklet)

(L) "TURN THE PAGE. FIRST IS A PICTURE OF A BOX AND A STAR." (point to first picture). "NEXT IS A PICTURE OF A STAR AND A STAR. PUT AN 'X' ON THE ONE I SAY: BOX, STAR."

(P) Make sure they all do it correctly and understand the instructions.

5. (L) "TURN TO THE NEXT PAGE. NOW WE ARE GOING TO LISTEN TO THE TAPE RECORDER. EVERYONE PLEASE BE VERY QUIET AND LISTEN, AND DO EXACTLY WHAT THE TAPE TELLS YOU TO DO."

Turn on recorder.

6. (P) As recorder plays, walk around among the students, checking to see that they are on the right page, etc. If a page is missing, or there are two the same, tell them not to worry, and help them go on to the next page, as the recorder gets to it.

7. (P) If there are any disturbances, such that you see some of them didn't hear the recorder, stop the tape, back it up, and tell them you are going to play the last one over so that everyone can hear it. Make sure the room is as quiet as possible at all times. If necessary, take out of the room anyone who continues to disturb the rest.

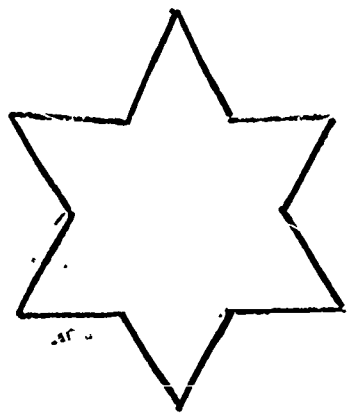
8. (P) At end of test, collect booklets, and mark on front of each booklet the teacher's name and the grade.

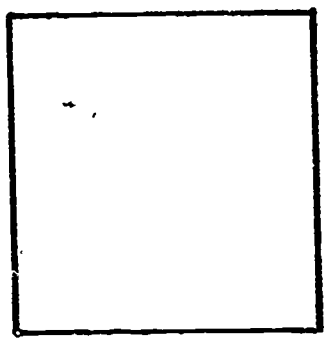
PAIRED ITEMS (test formats 3 & 4)

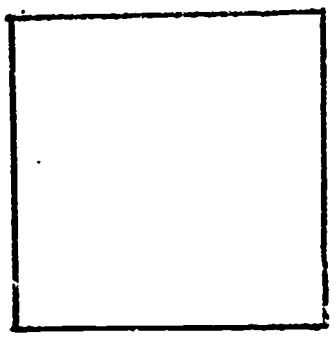
Sample items:	i.	star/box	gate/date
	ii.	box/star	star/star
Test:	1.	gate/gate	gate/date
	2.	bat/back	bat/bat
	3.	chop/chop	chop/shop
	4.	tug/tub	tub/tub
	5.	cot/cot	cot/pot
	6.	gum/gun	gun/gun
	7.	sack/sack	shack/sack
	8.	gas/gas	gas/gash
	9.	tool/tool	pool/tool
	10.	deer/deer	deer/beer
	11.	seat/feet	feet/feet
	12.	top/cop	cop/cop
	13.	watch/watch	wash/watch
	14.	goat/goat	goat/boat
	15.	robe/road	robe/robe
	16.	fire/fire	fire/file
	17.	lock/lock	lock/rock
	18.	cap/cat	cat/cat
	19.	lass/laugh	laugh/laugh
	20.	tool/pool	pool/pool
	21.	bug/bug	bud/bug
	22.	gum/gum	gun/gum
	23.	cake/cape	cake/cake
	24.	beer/deer	beer/beer
	25.	back/back	back/bat

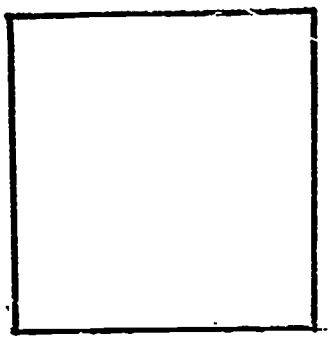
26.	mail/nail	nail/nail
27.	lass/lass	lass/laugh
28.	rock/lock	rock/rock
29.	boat/goat	boat/boat
30.	tug/tug	tub/tug
31.	pot/cot	pot/pot
32.	cape/cape	cape/cake
33.	top/top	cop/top

i.

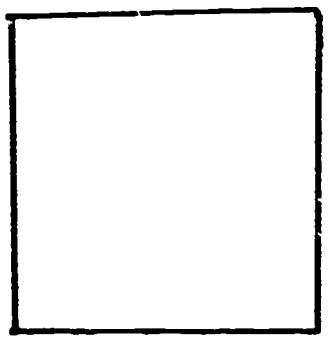


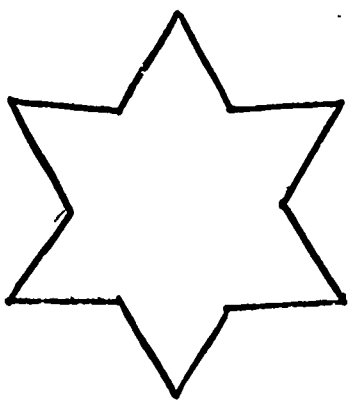


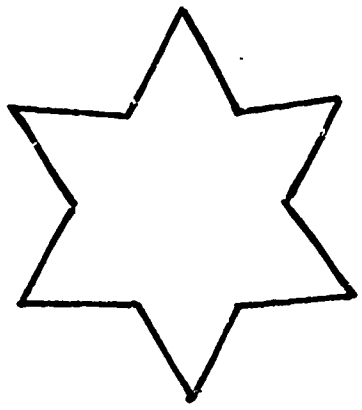


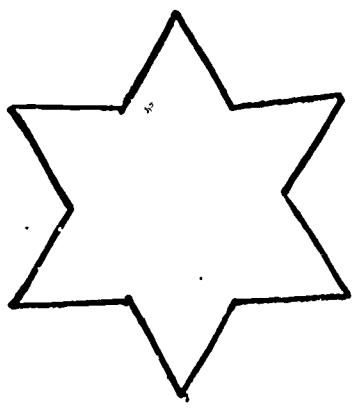


ii

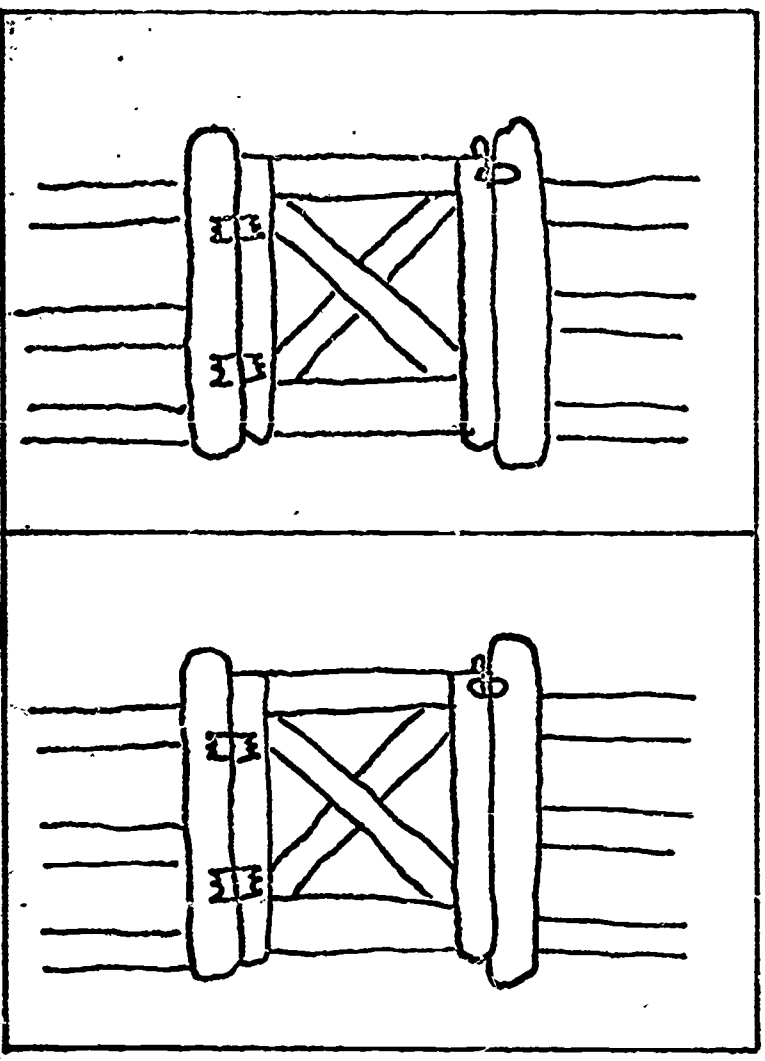


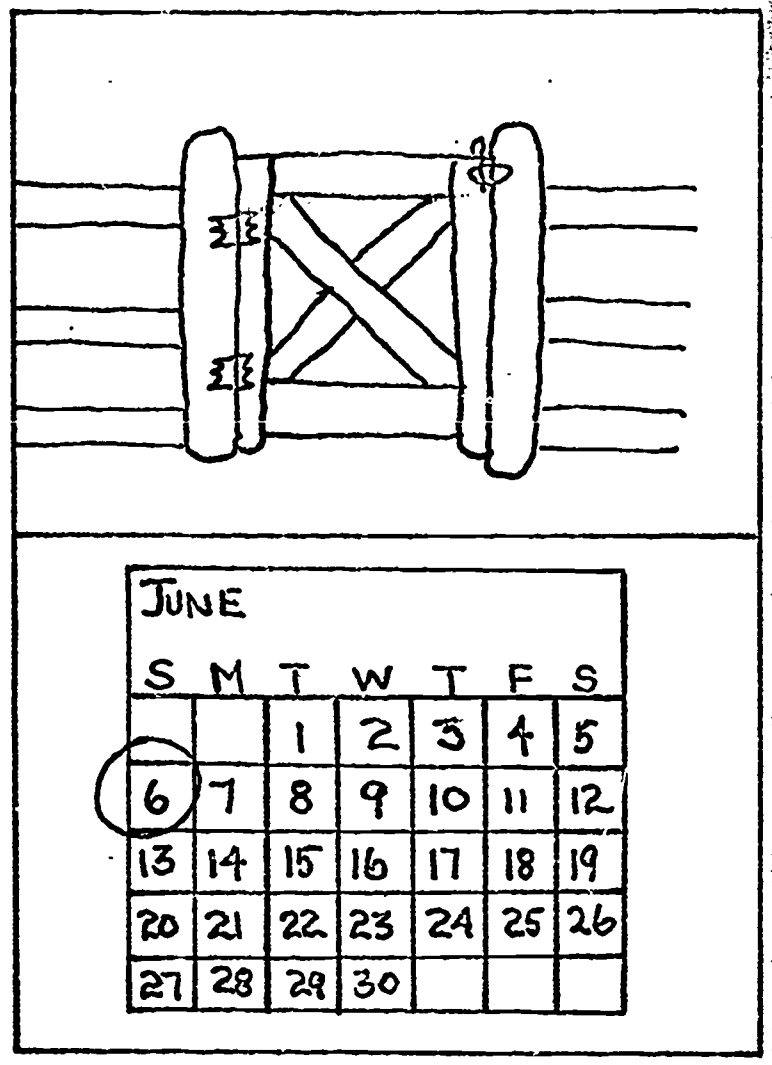






1

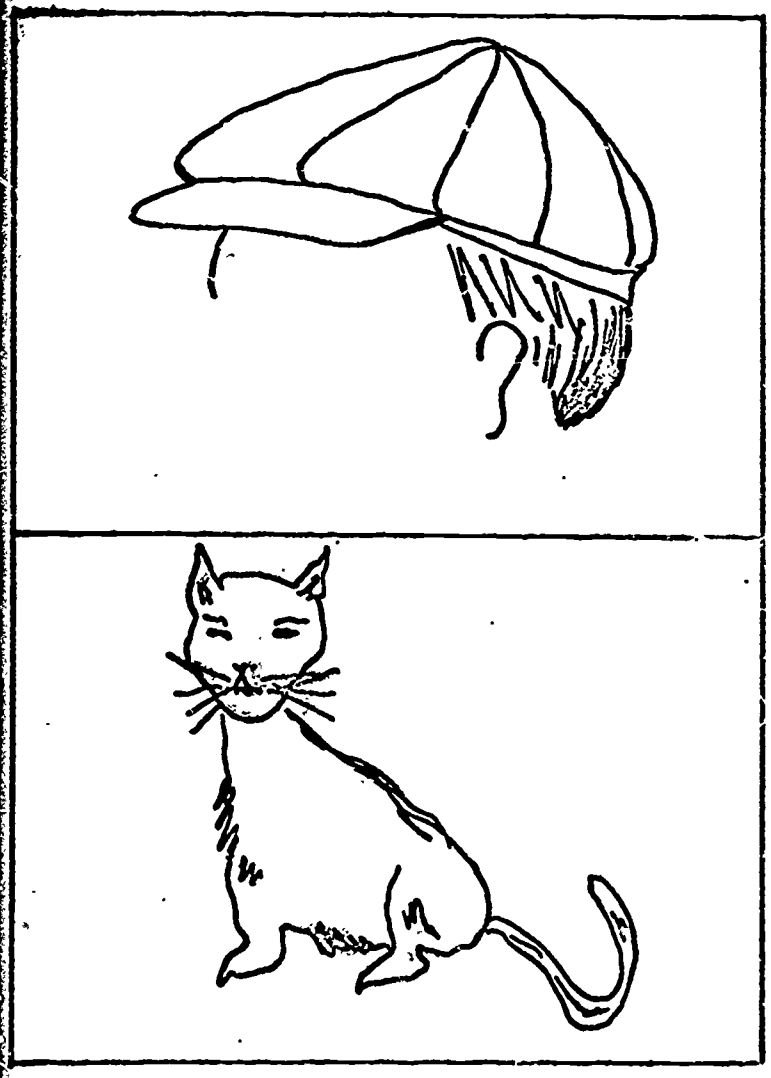


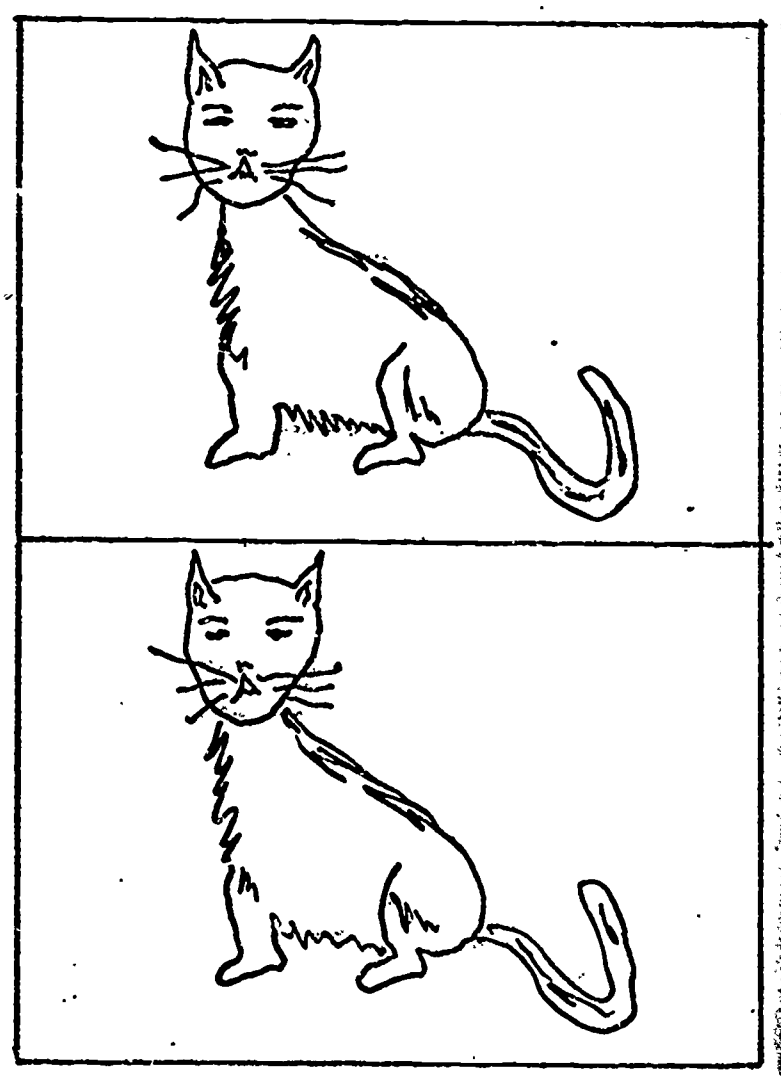


JUNE

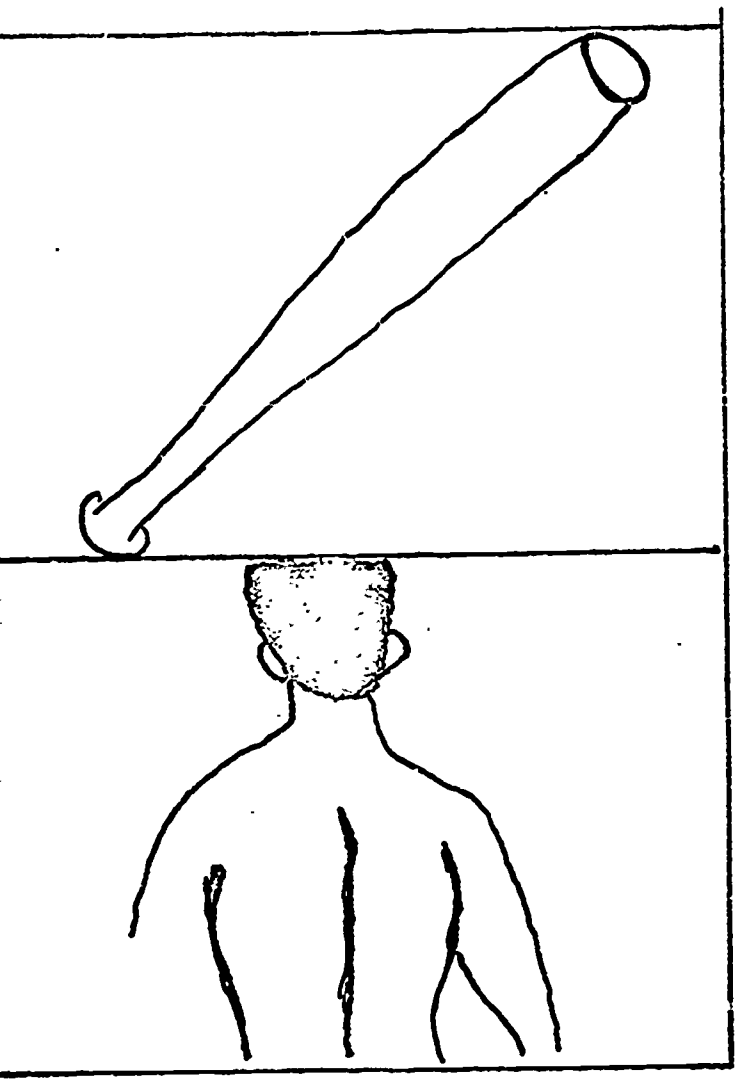
S	M	T	W	T	F	S
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30			

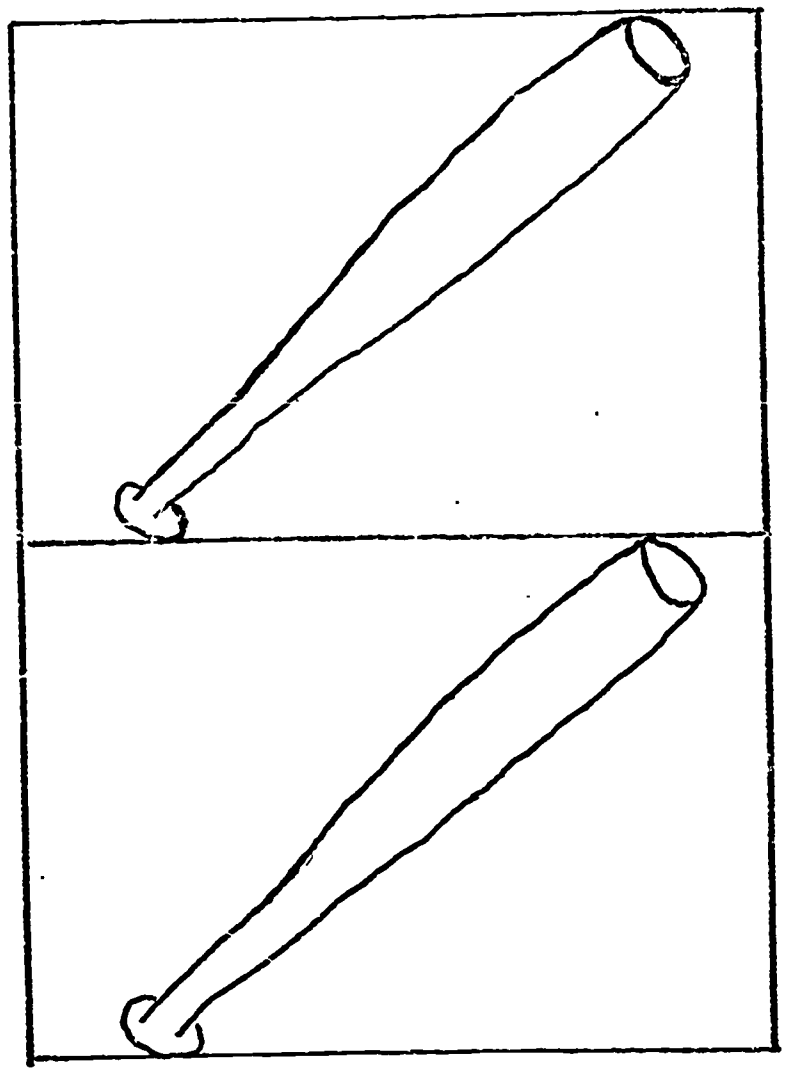
18



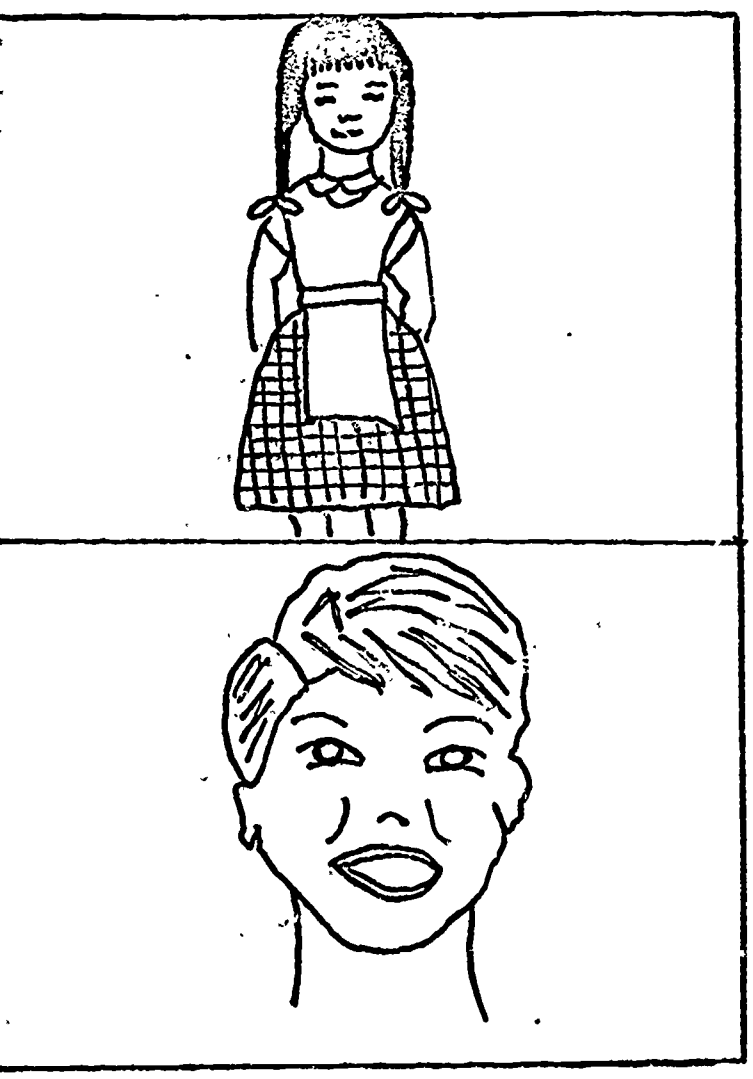


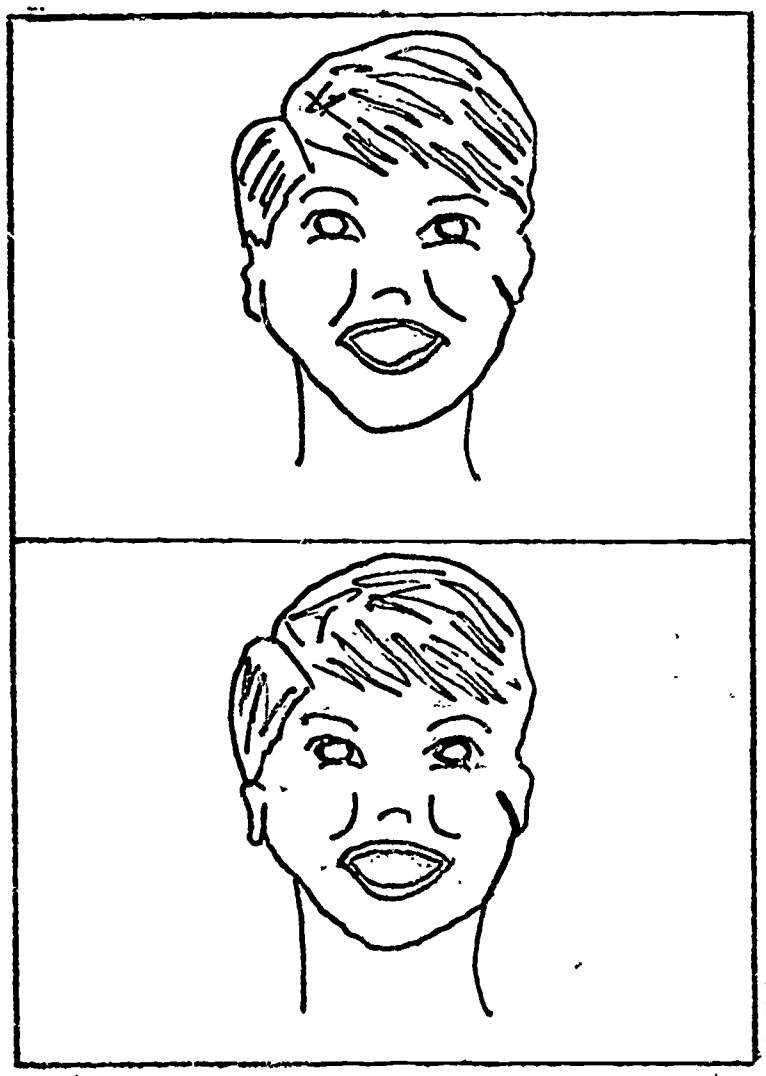
2



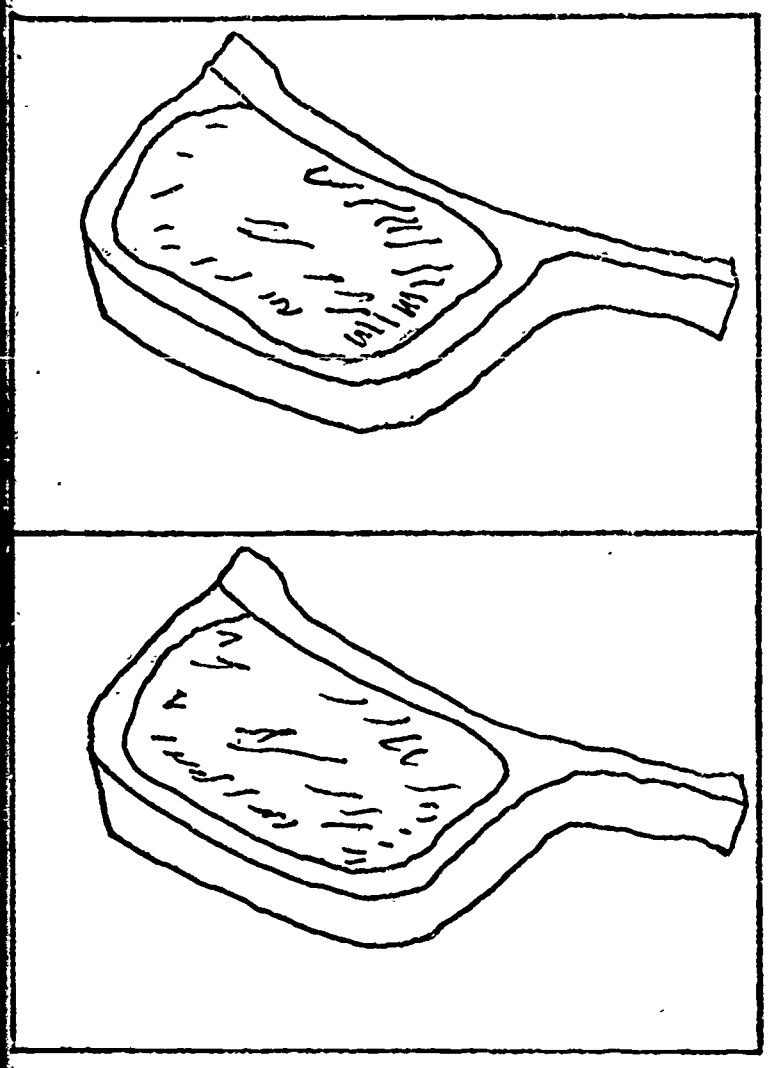


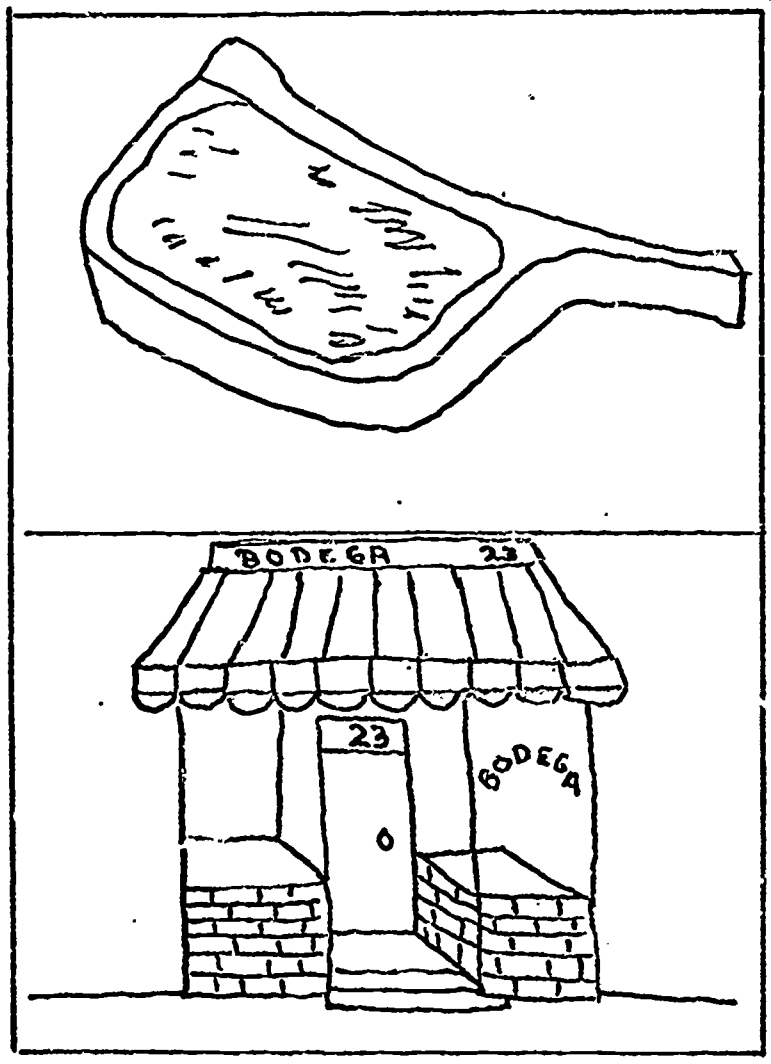
19



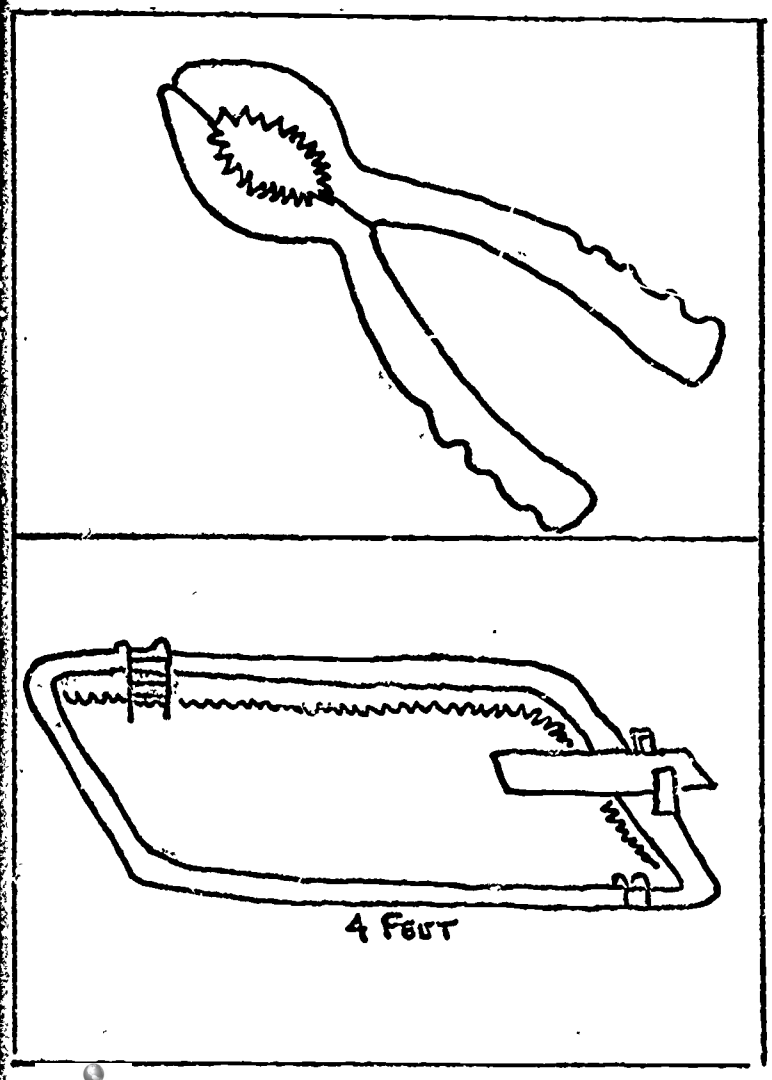


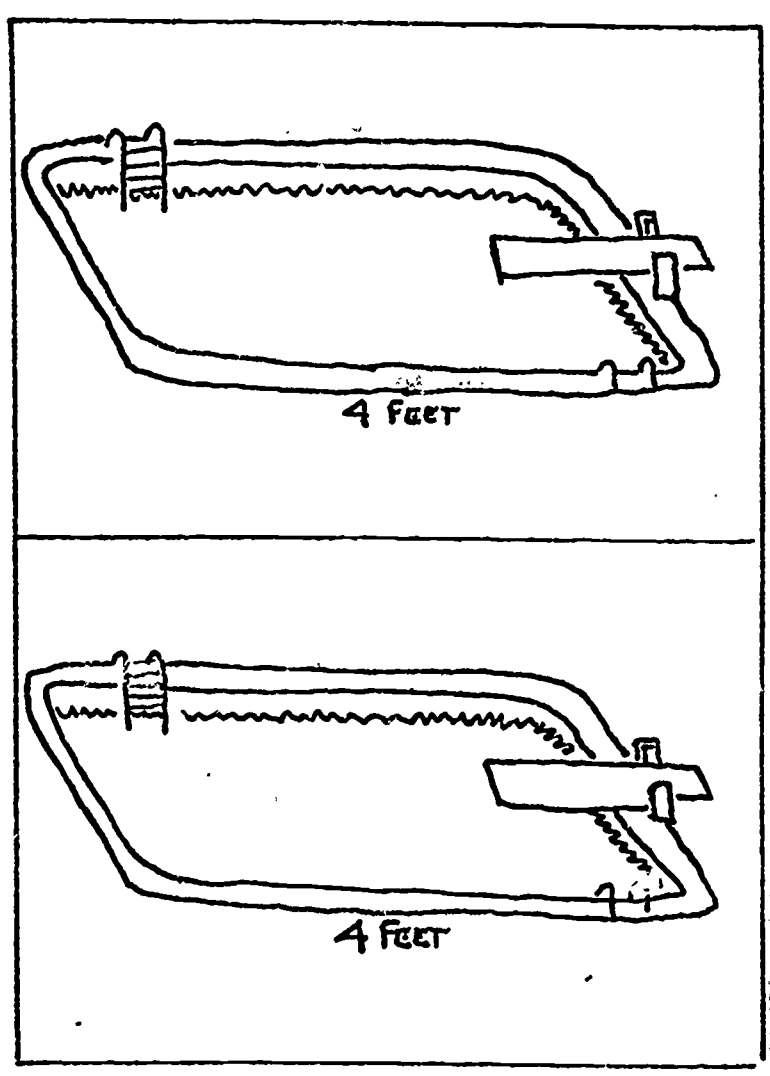
3



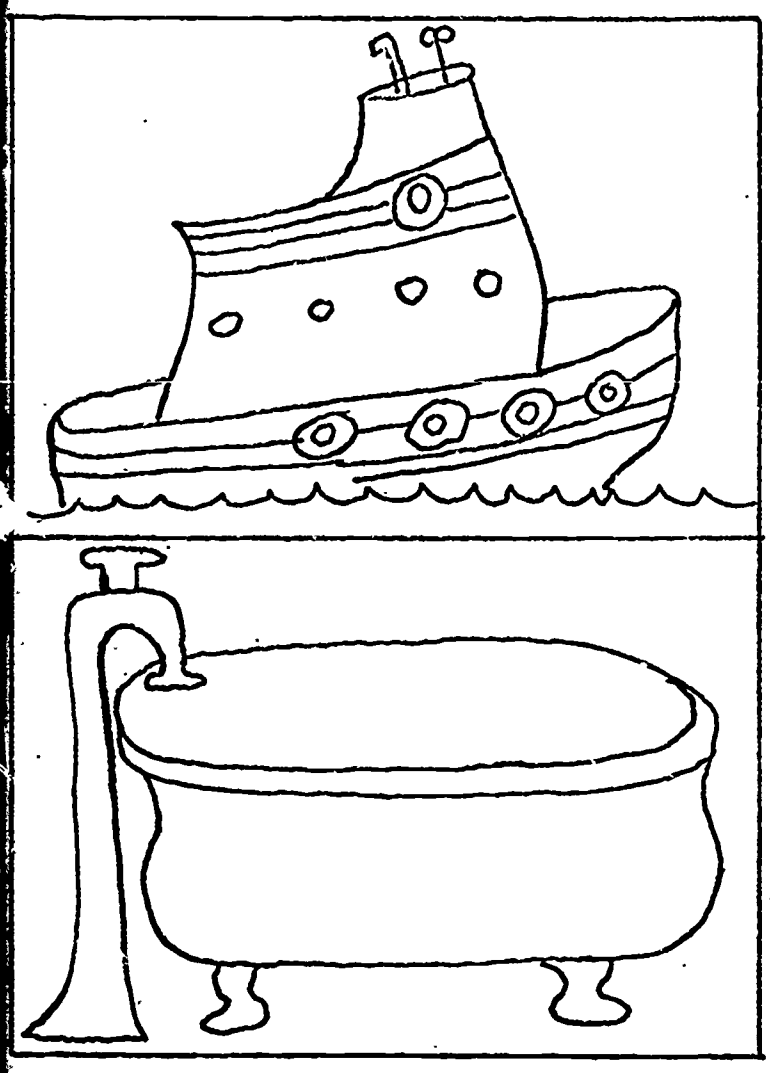


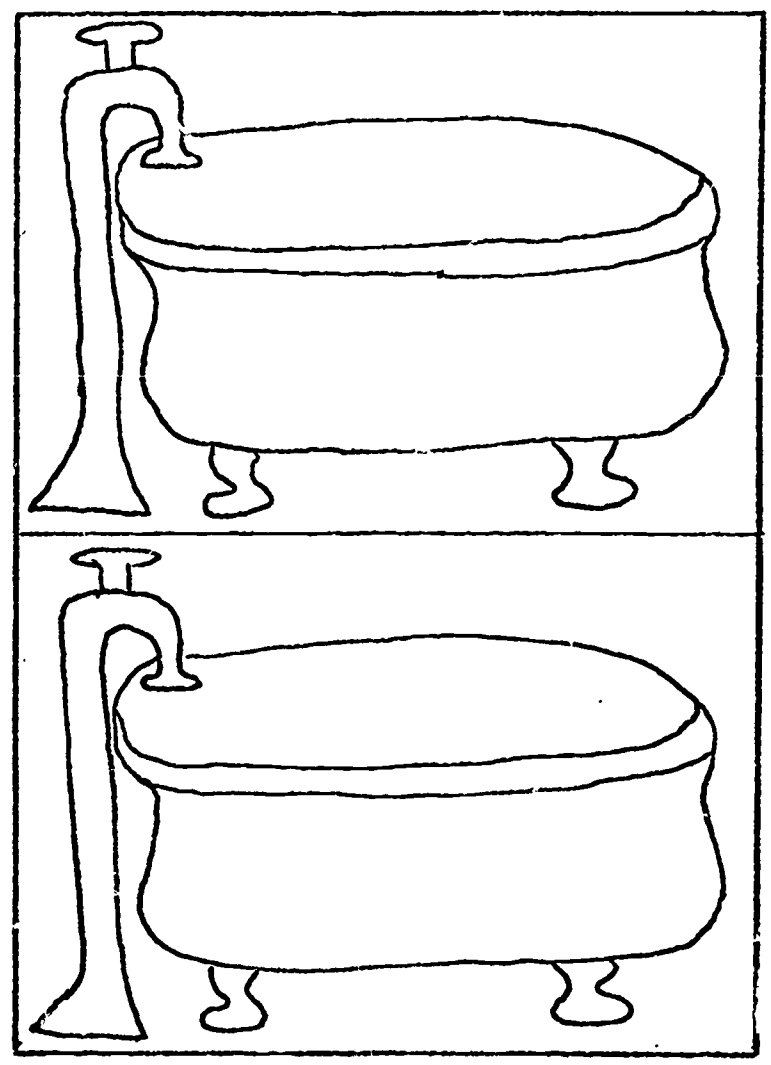
20



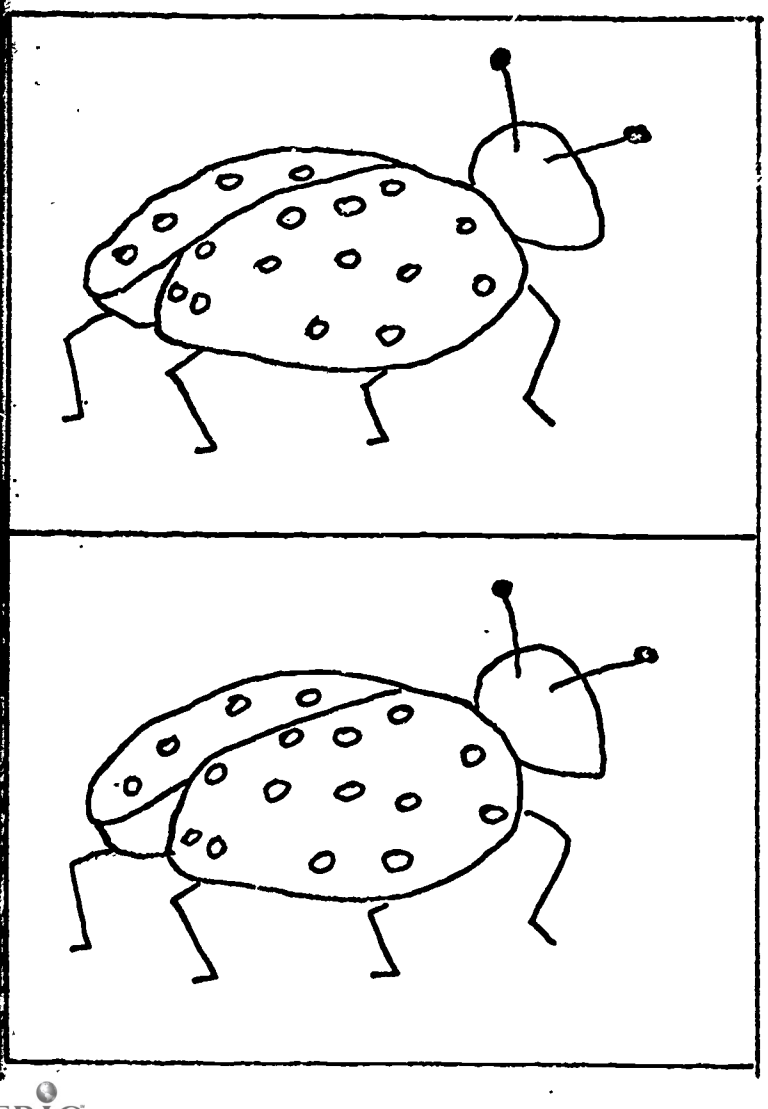


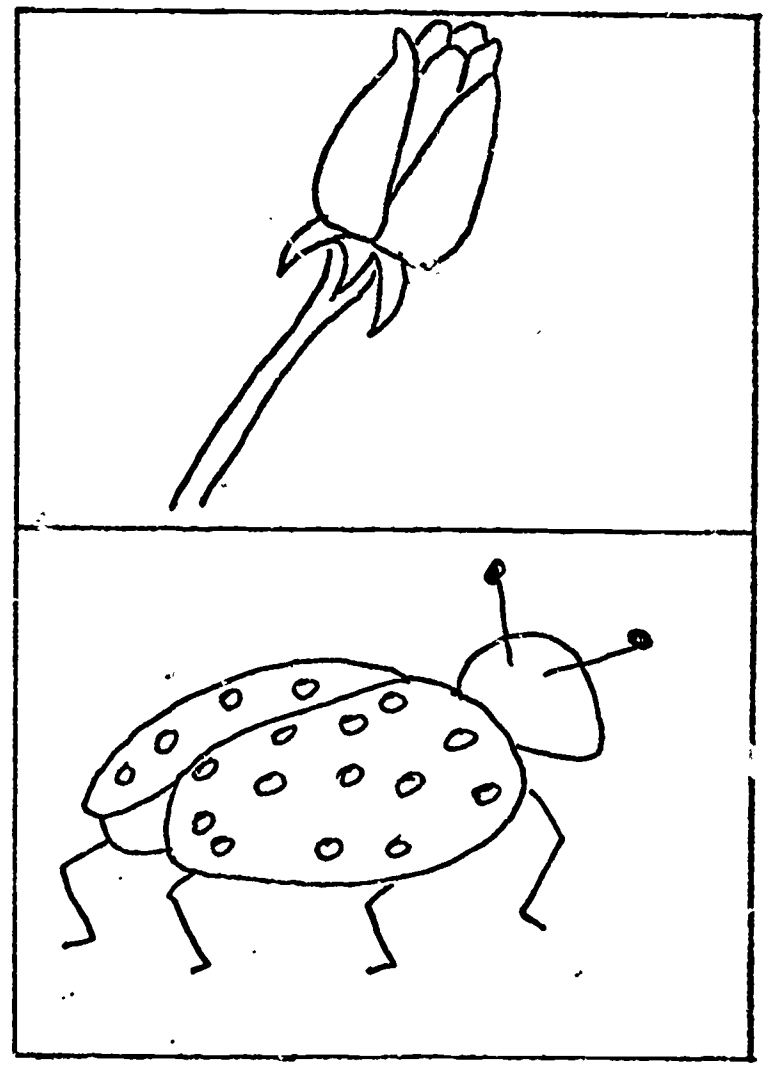
4



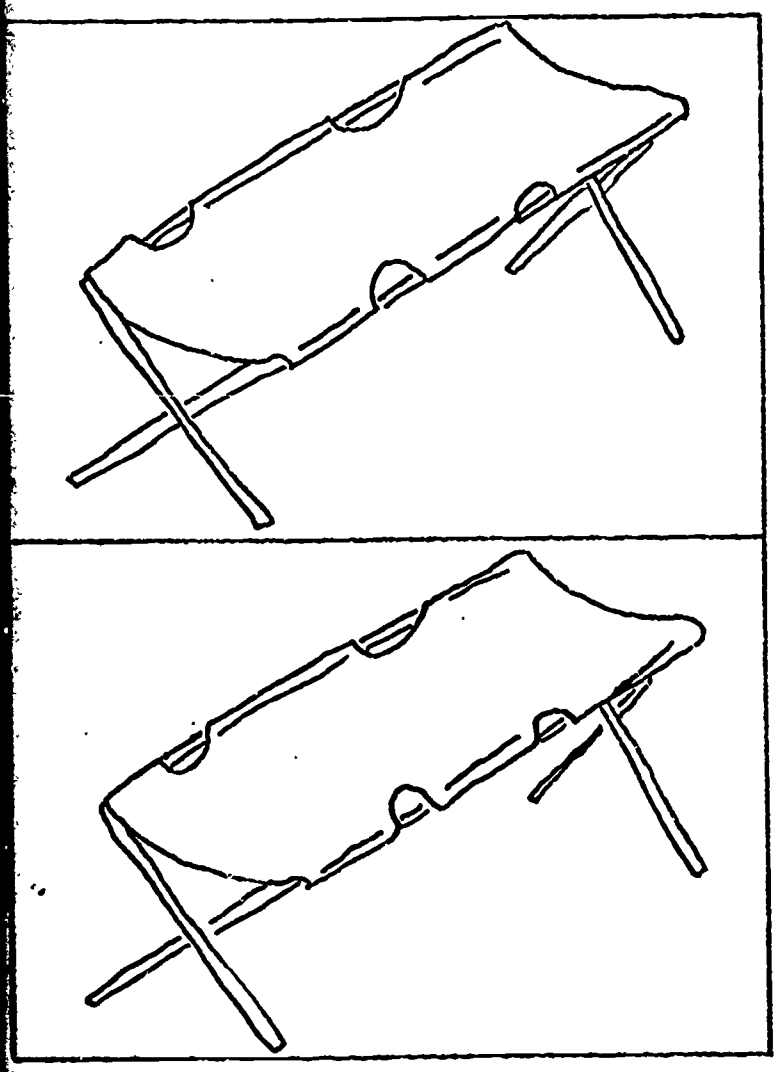


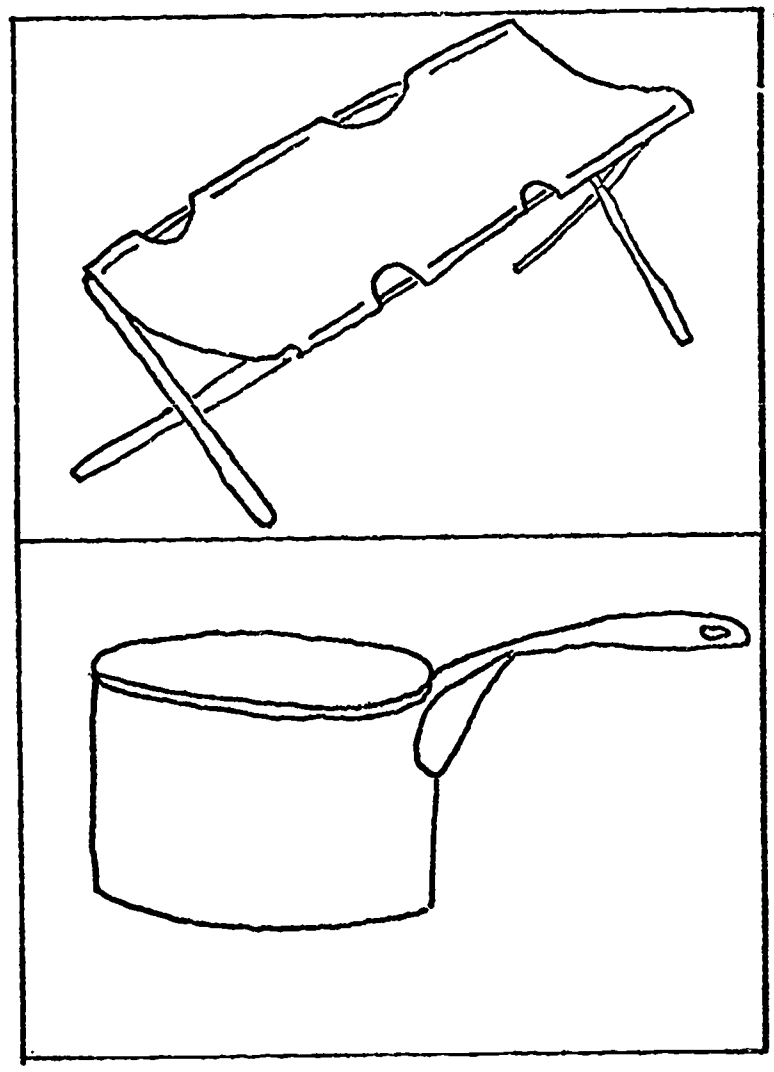
21



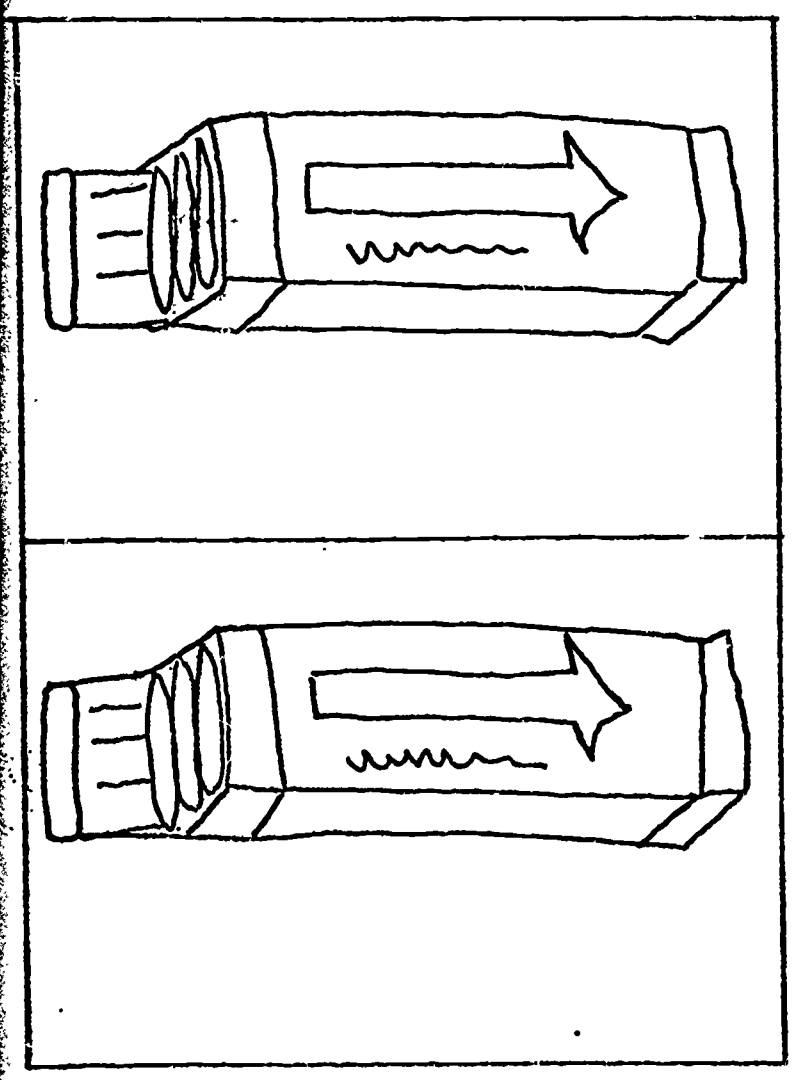


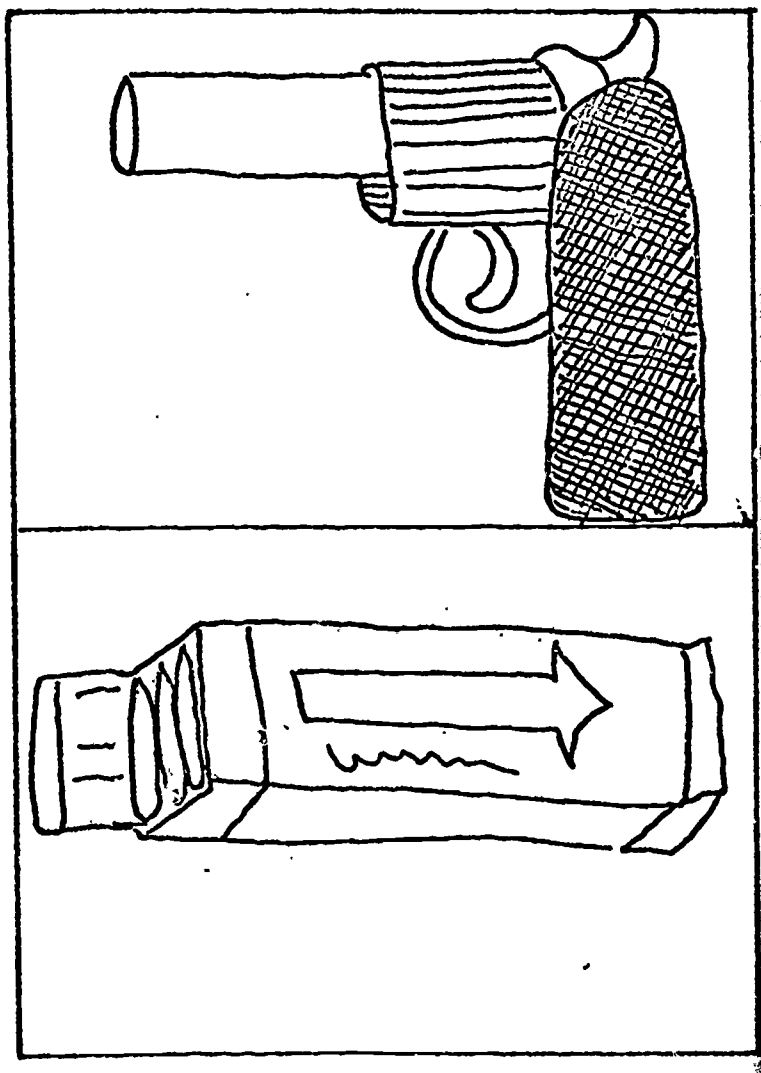
5



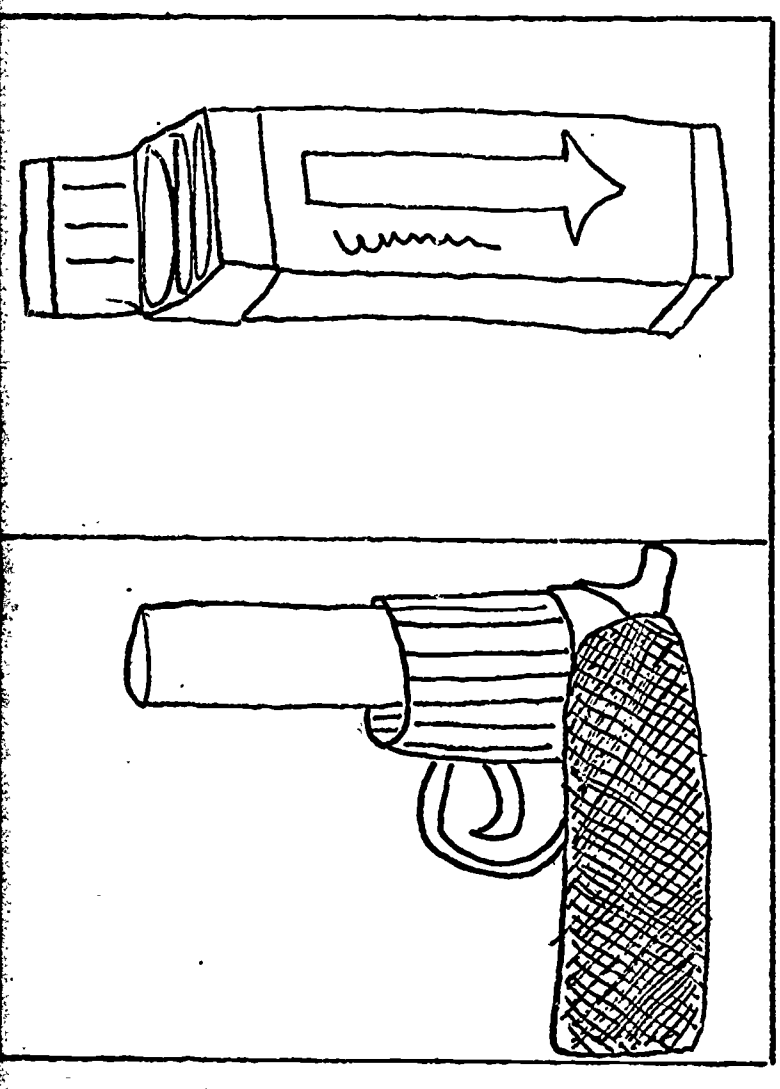


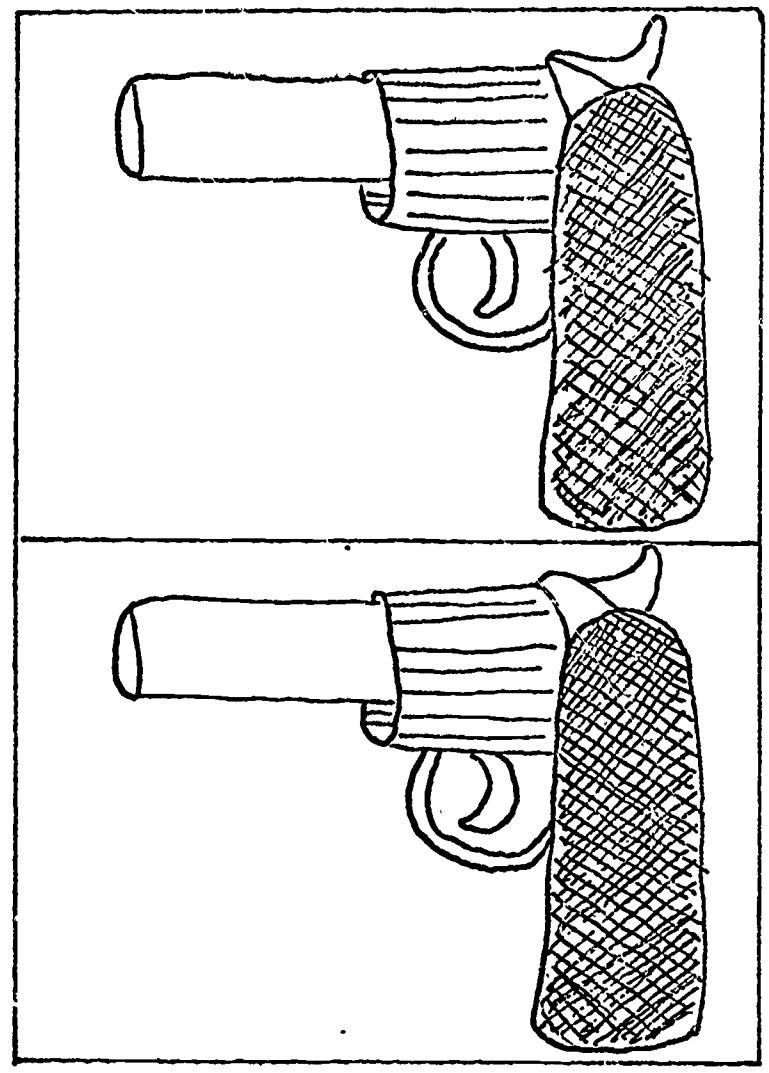
22



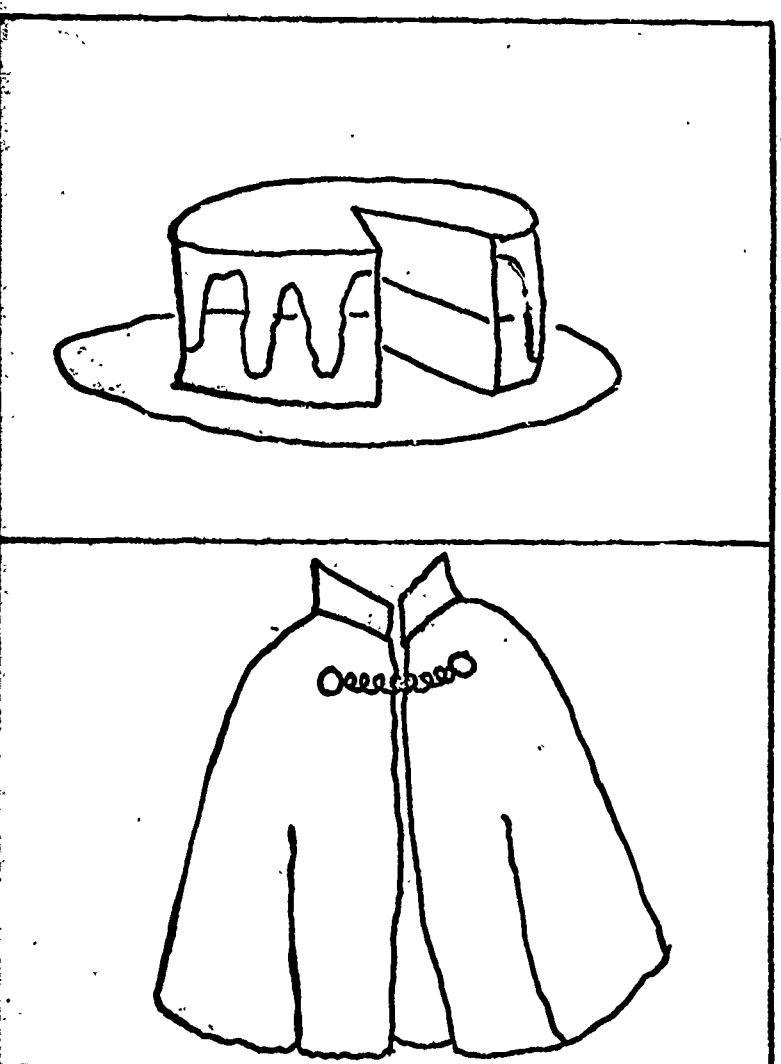


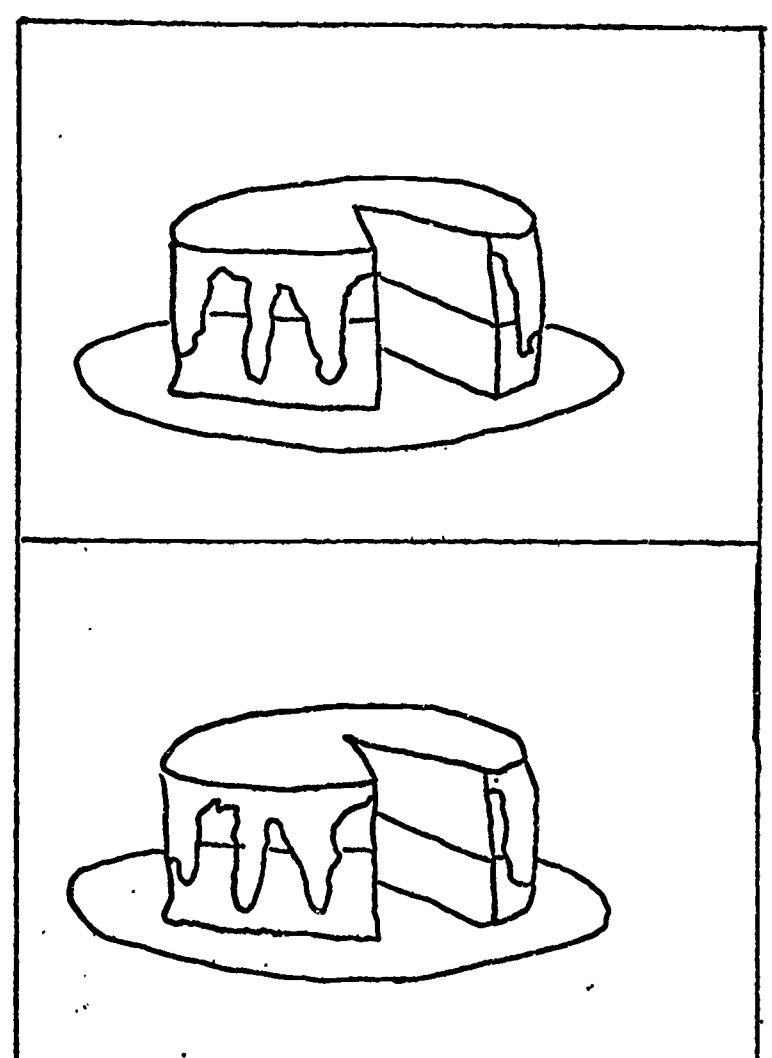
6



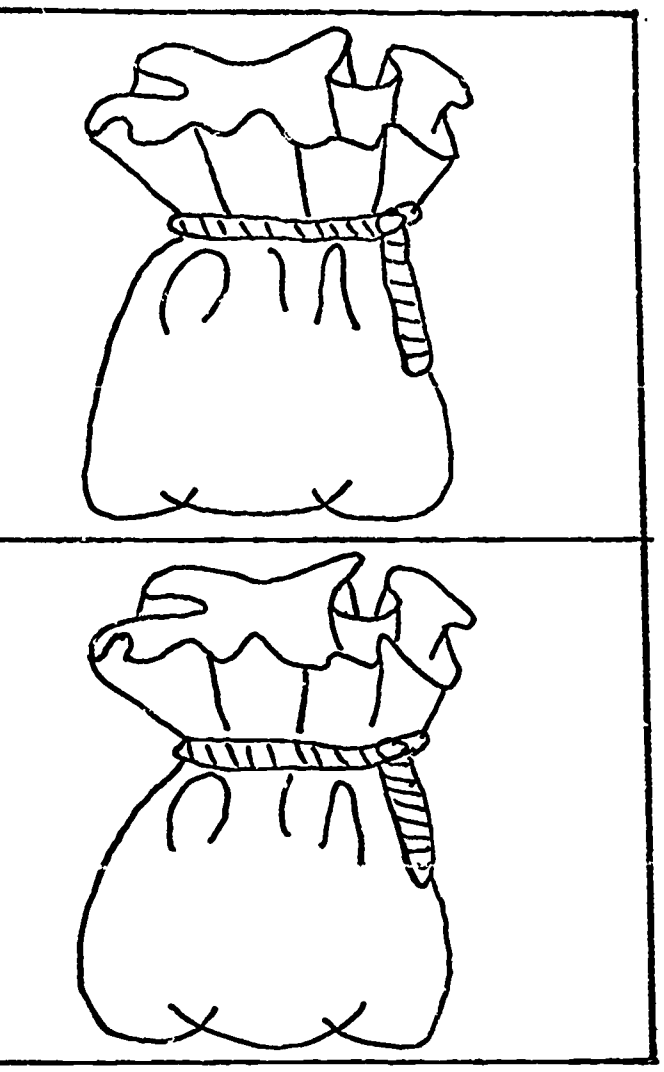


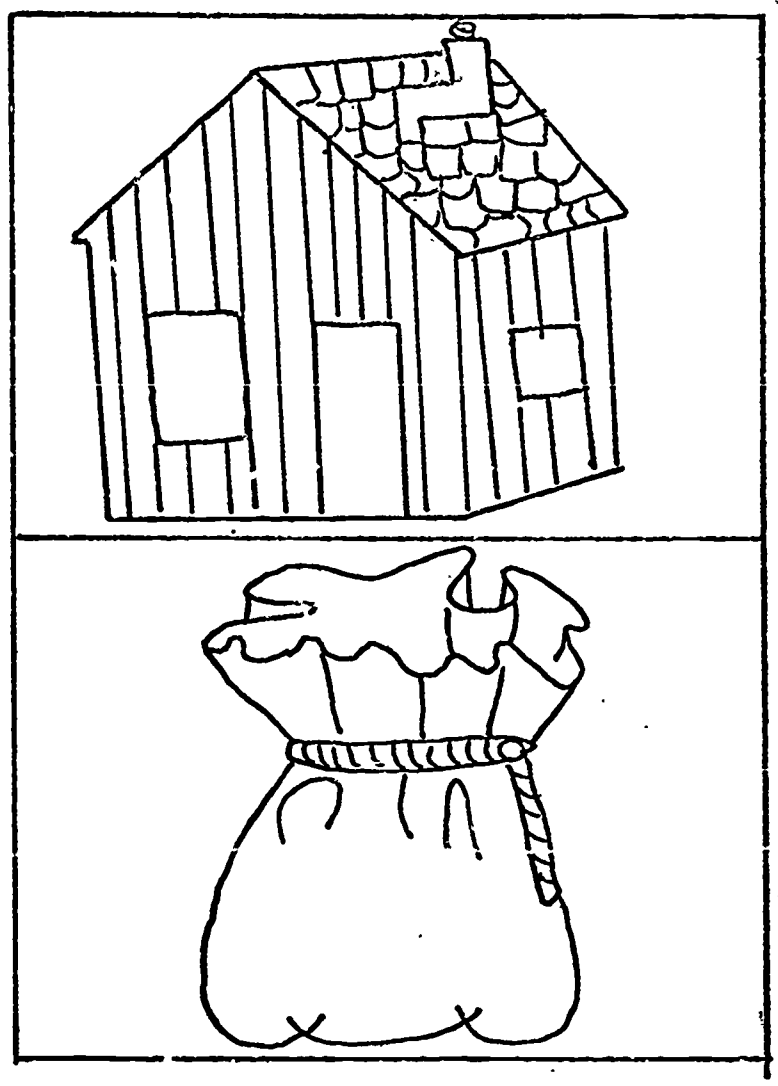
23



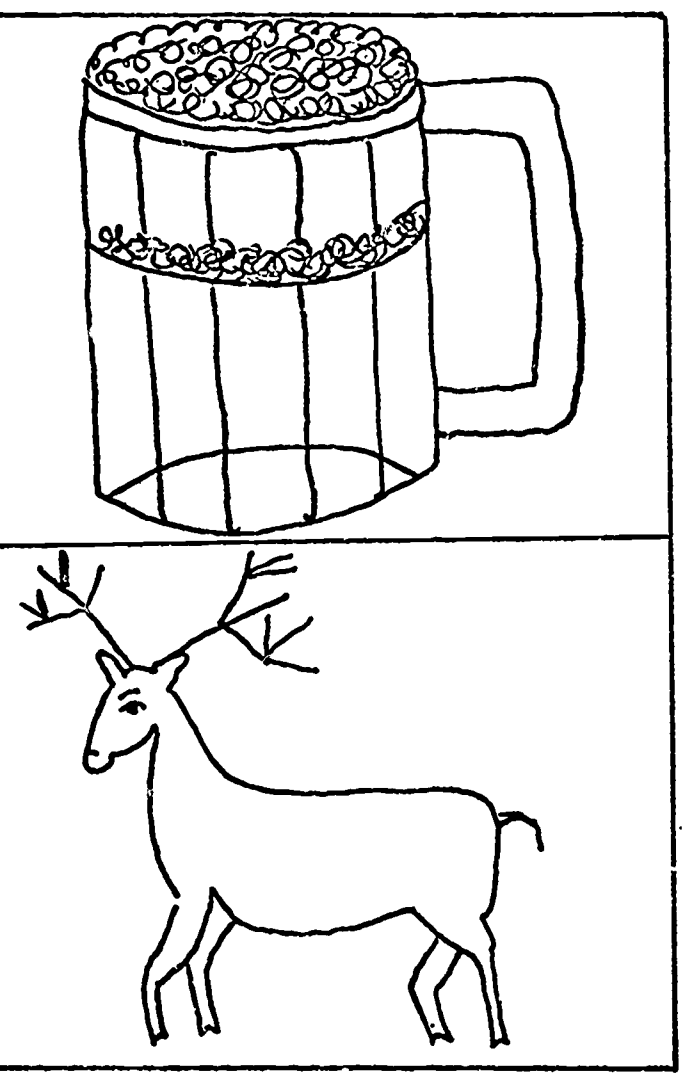


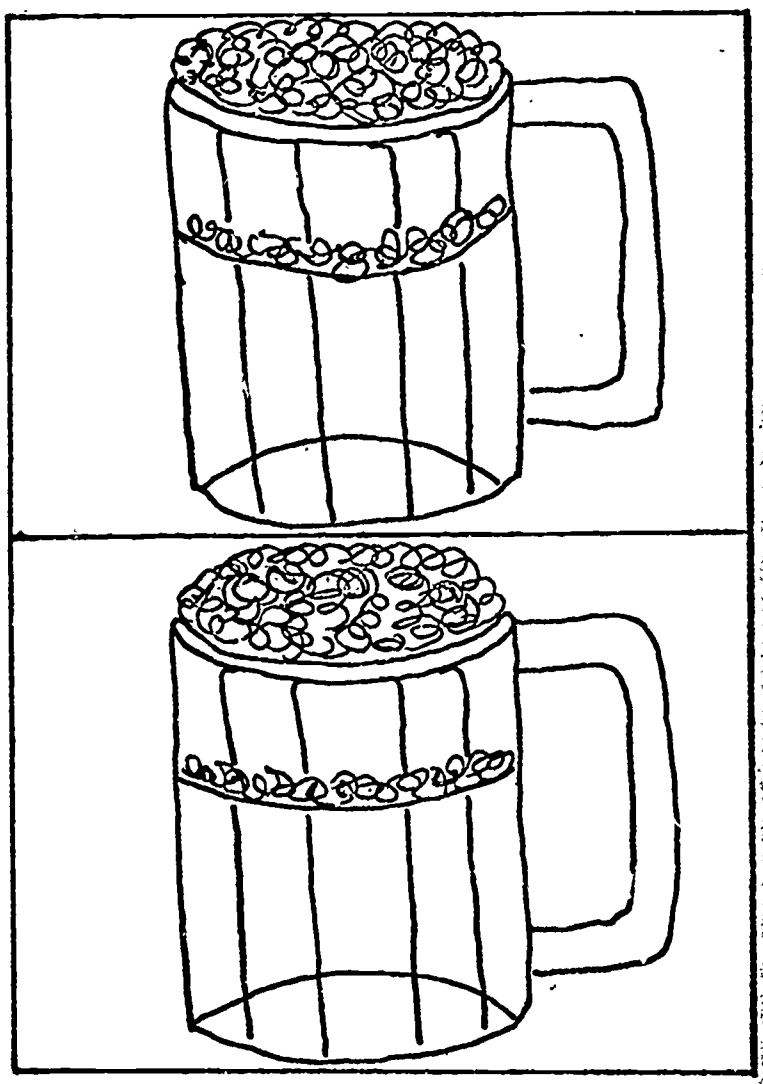
7



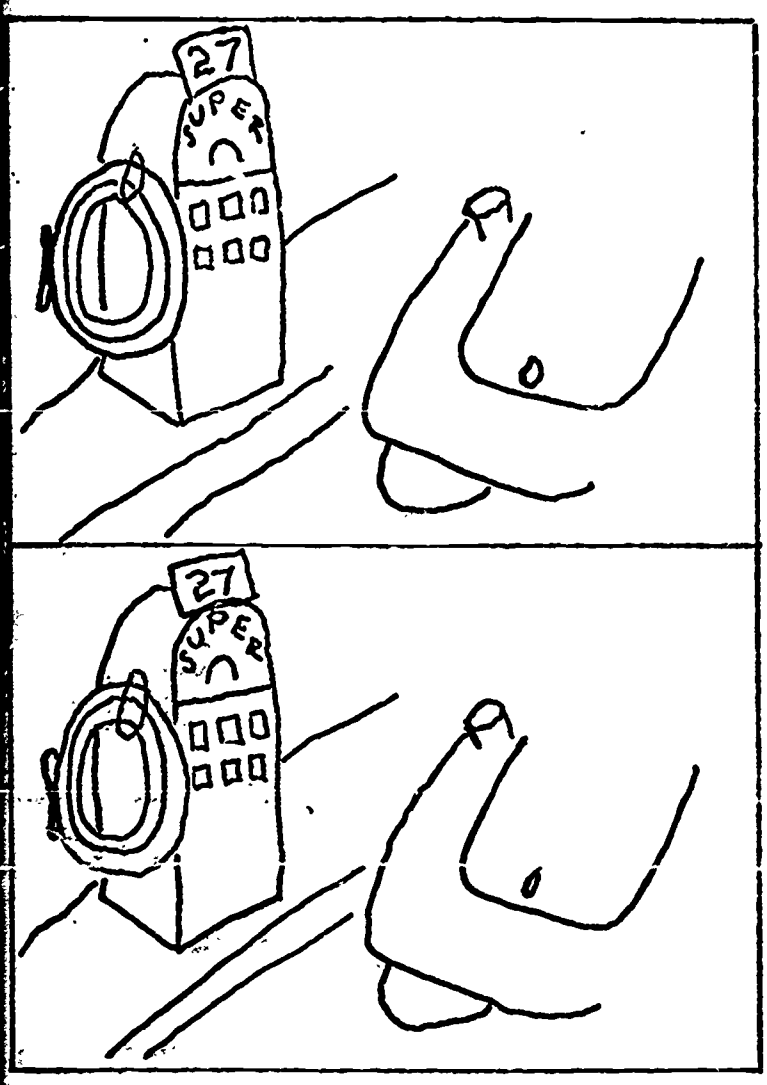


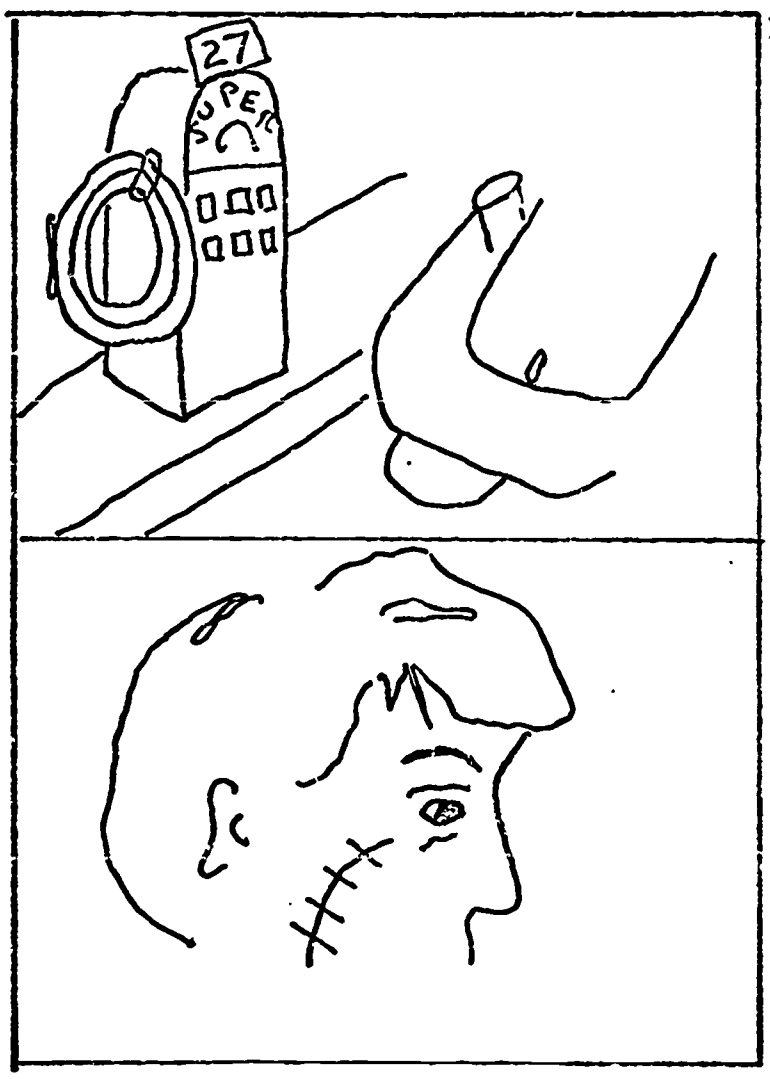
24



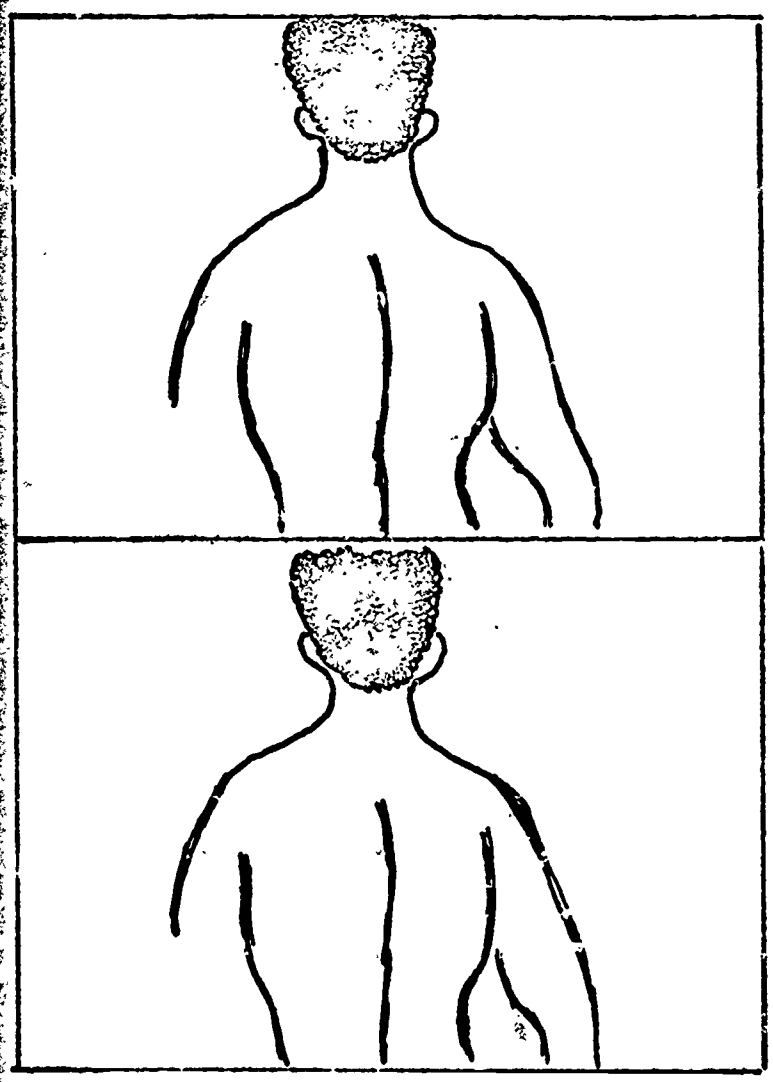


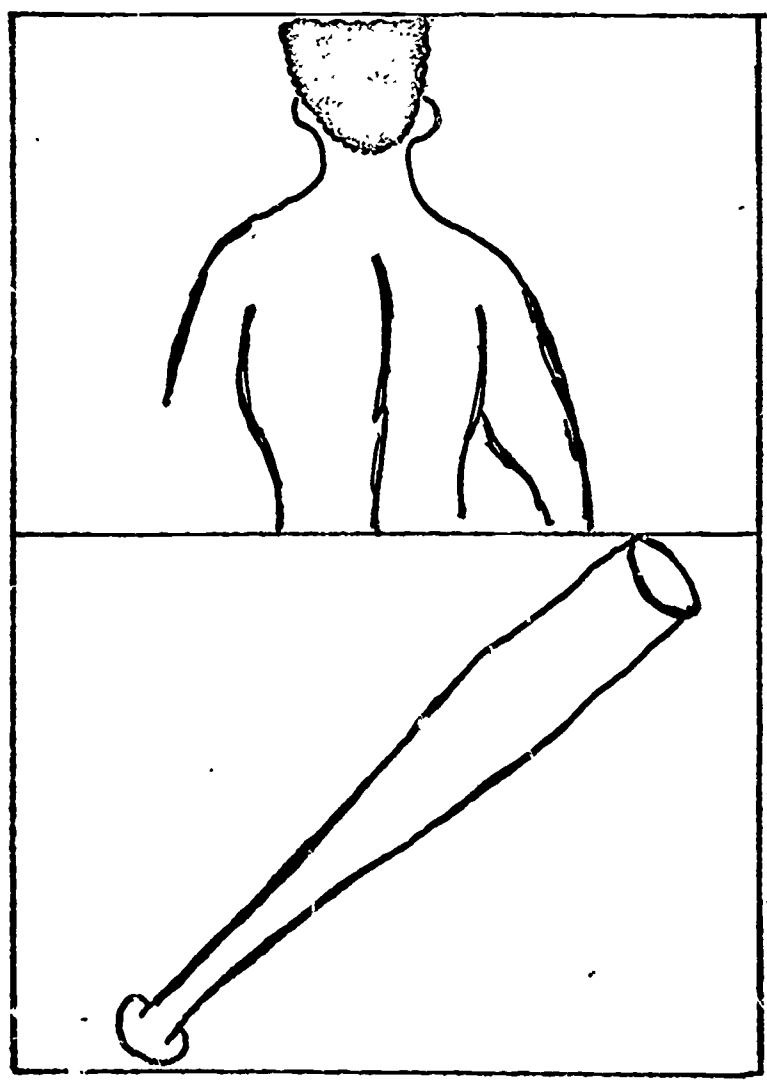
8



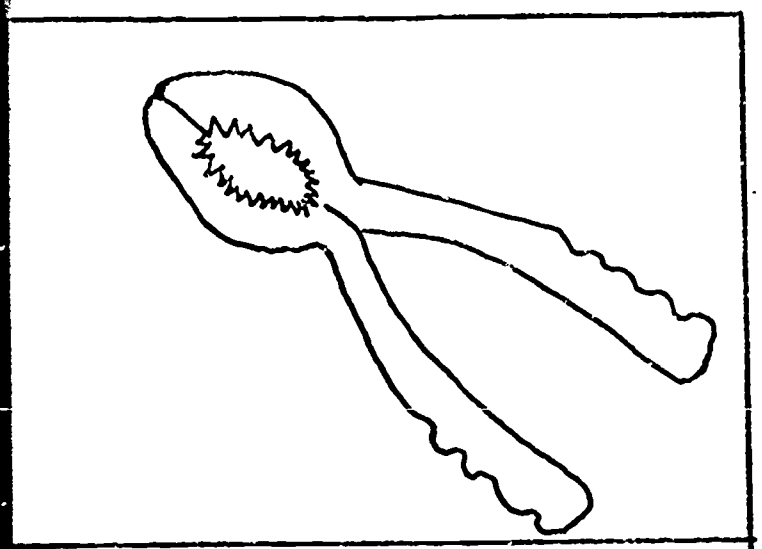


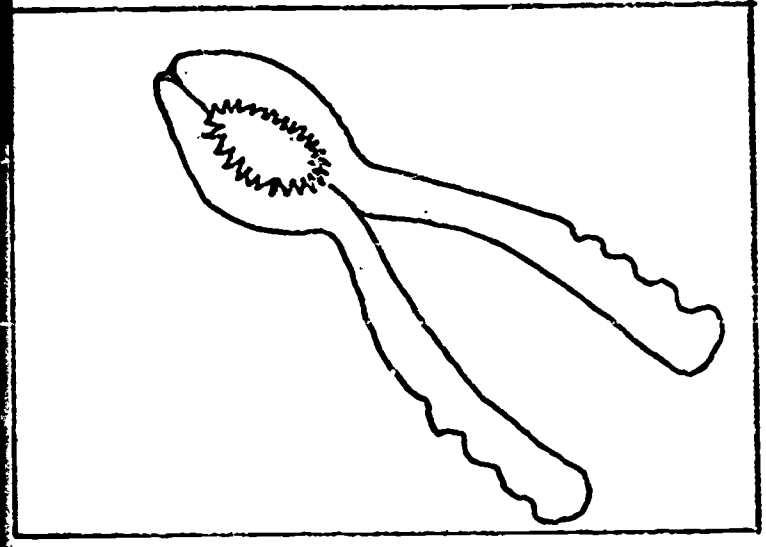
25

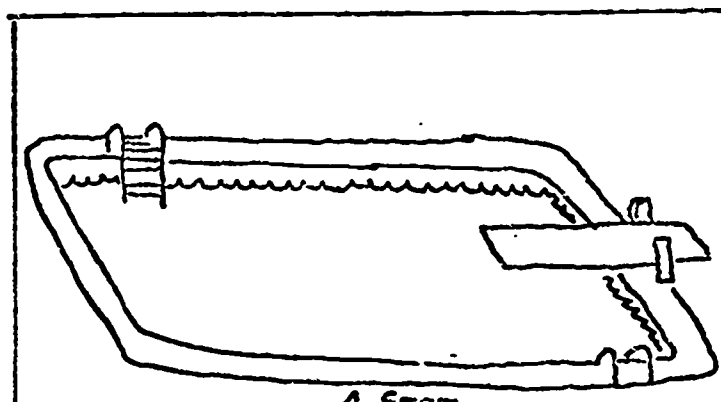


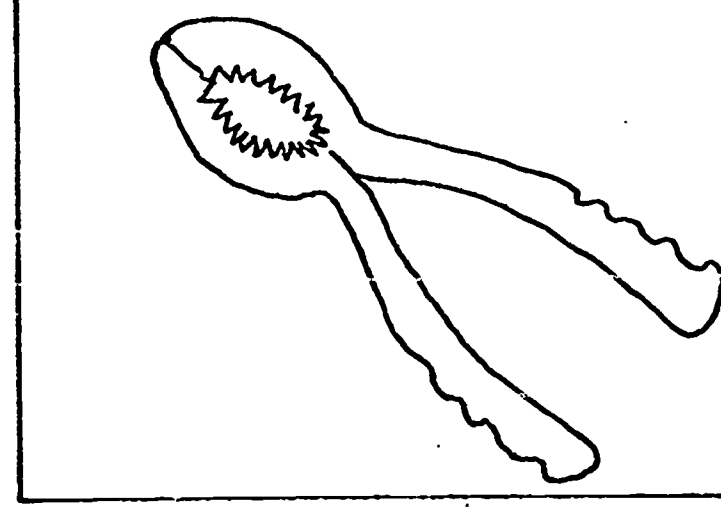


4

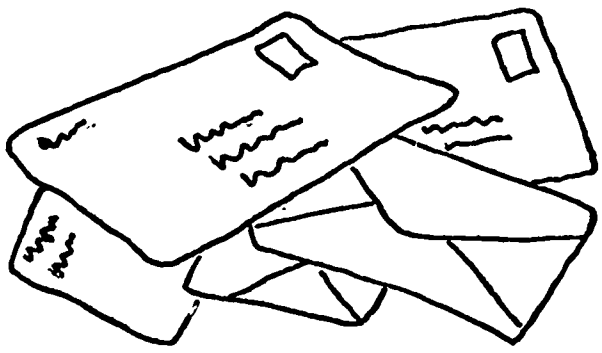


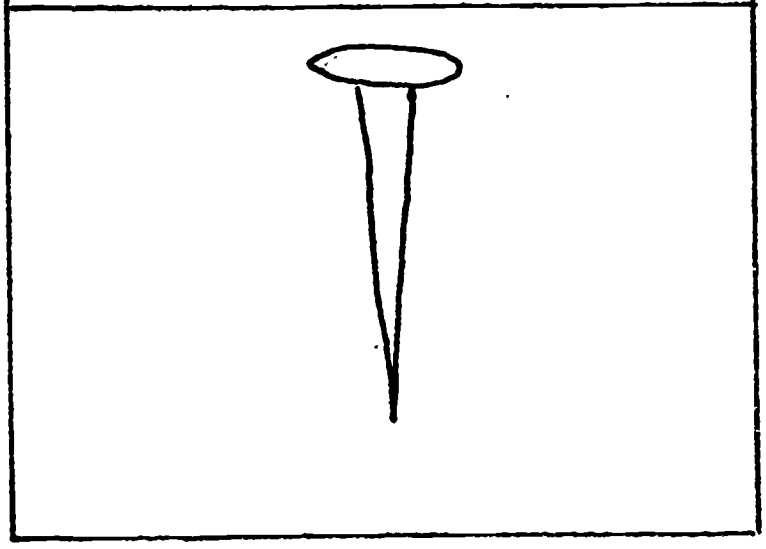


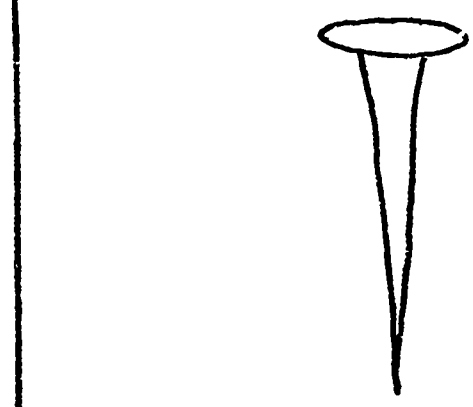


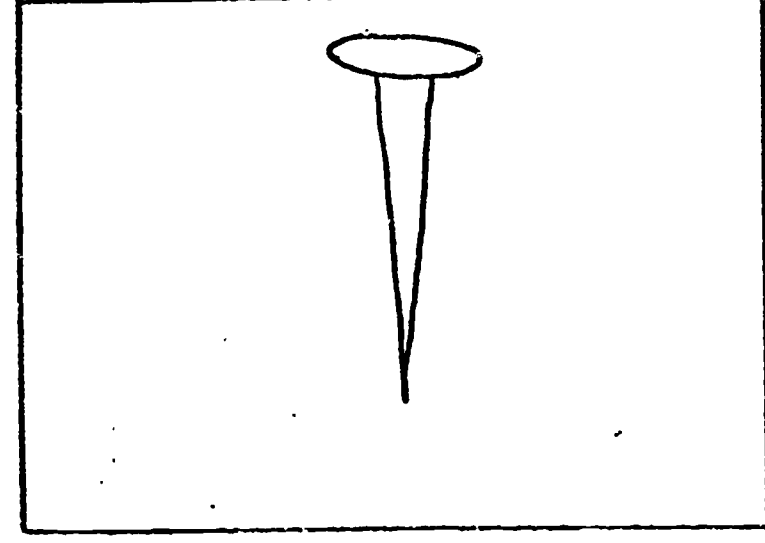


2b

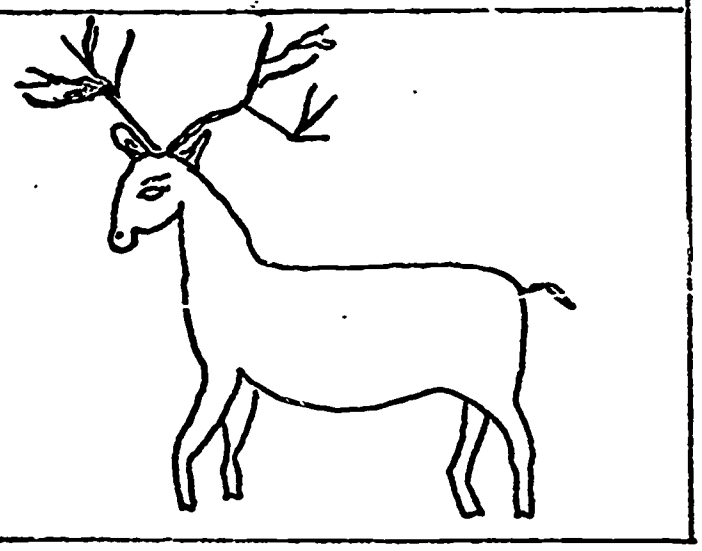


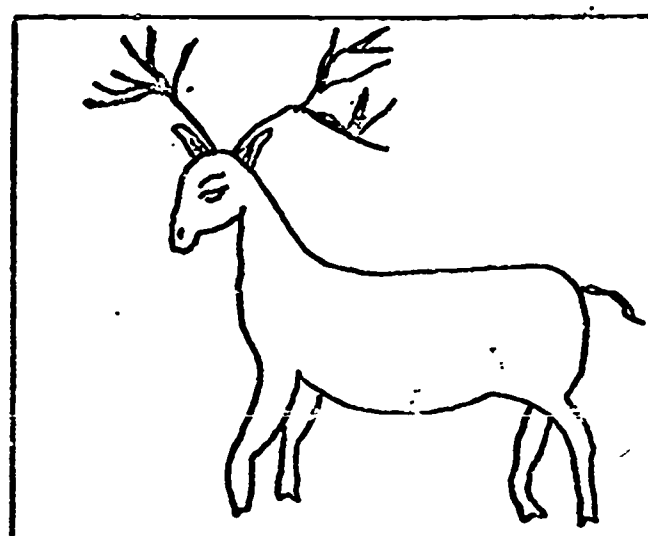


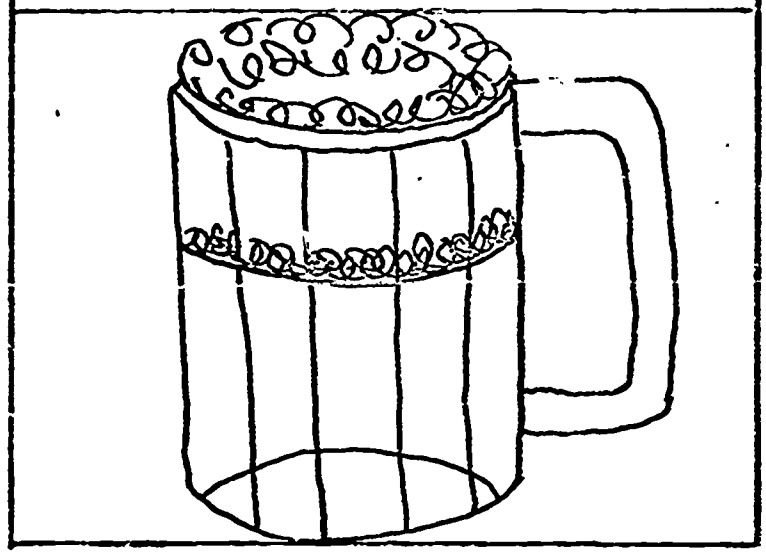




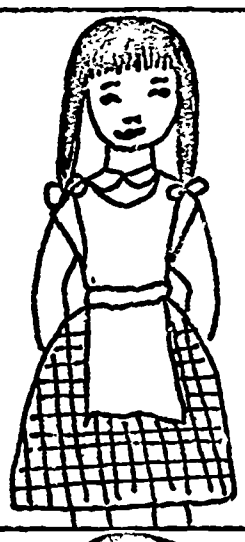
10

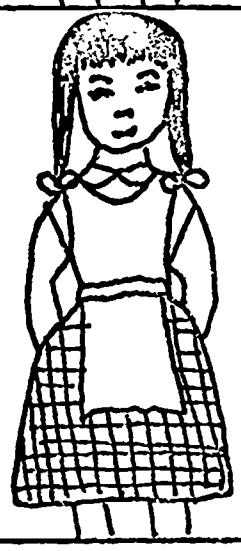


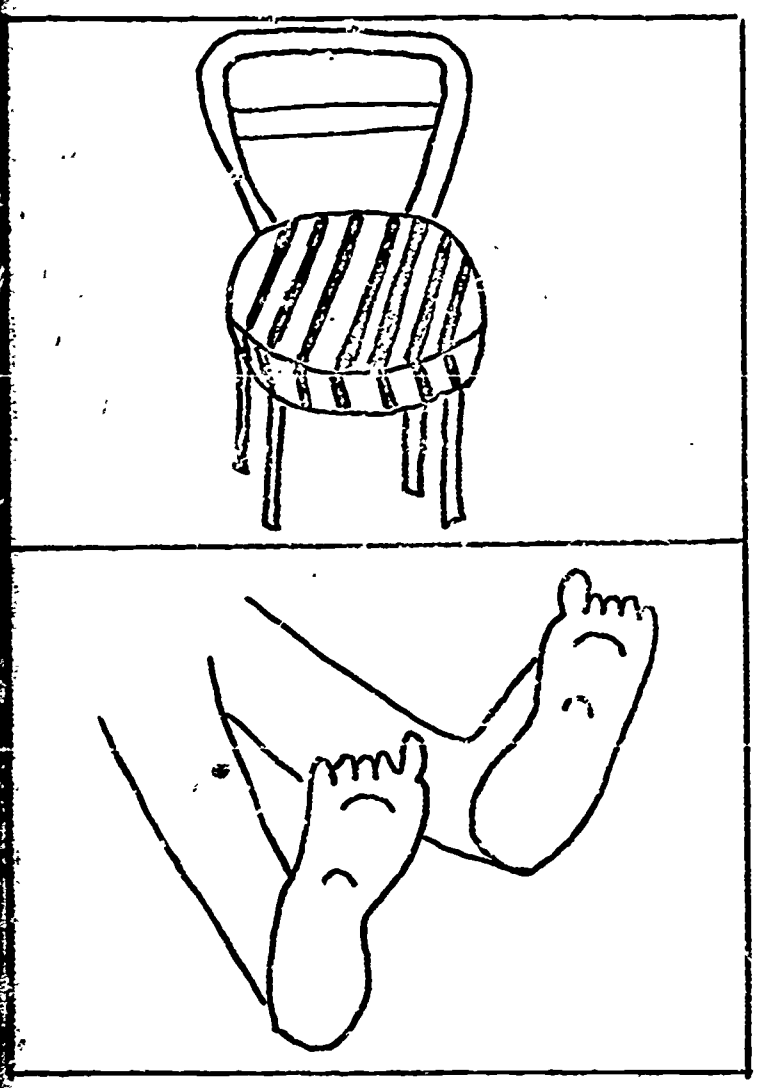


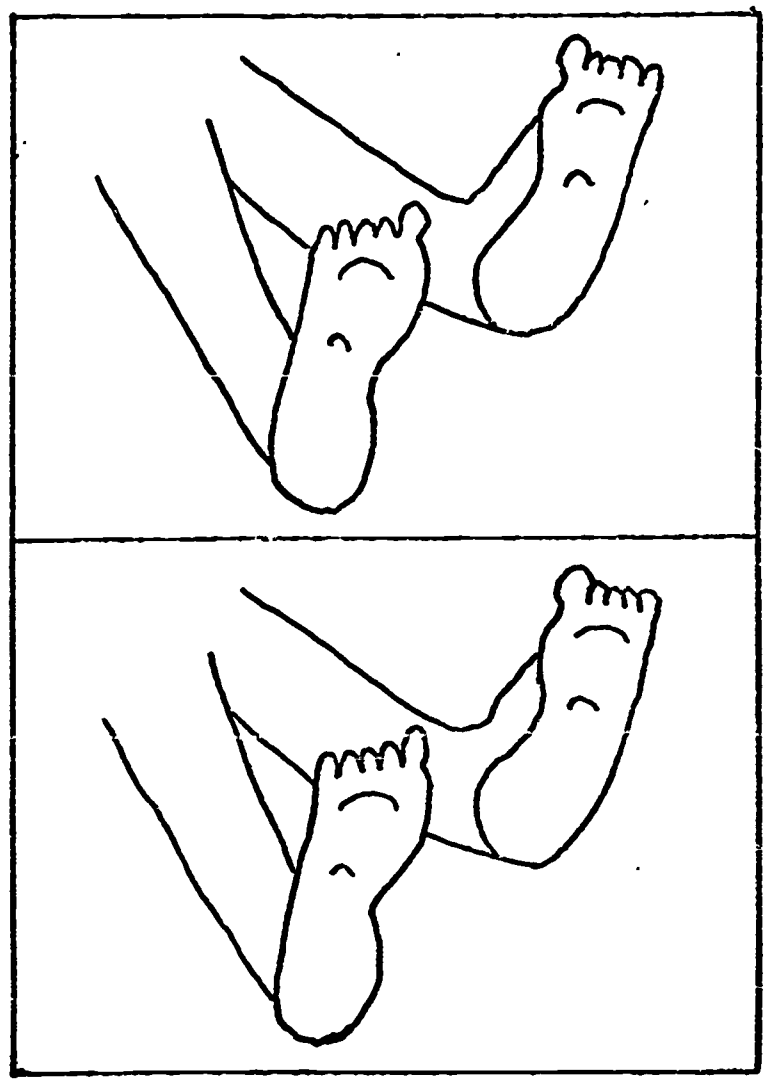


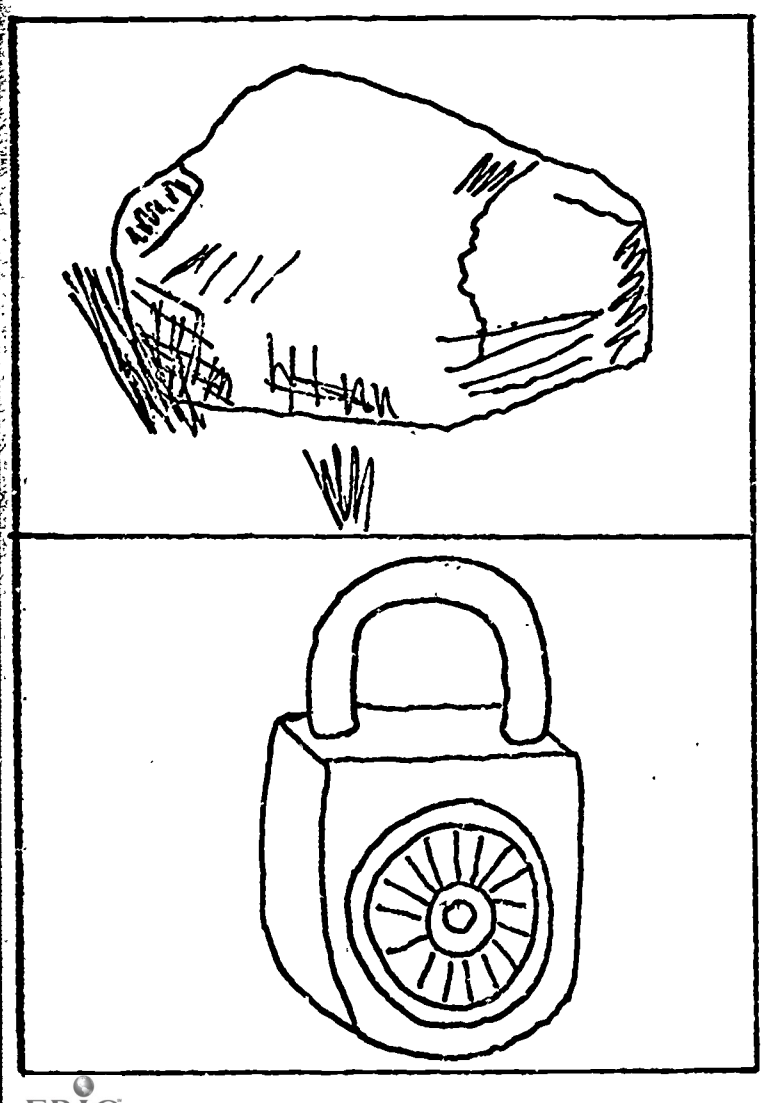
27

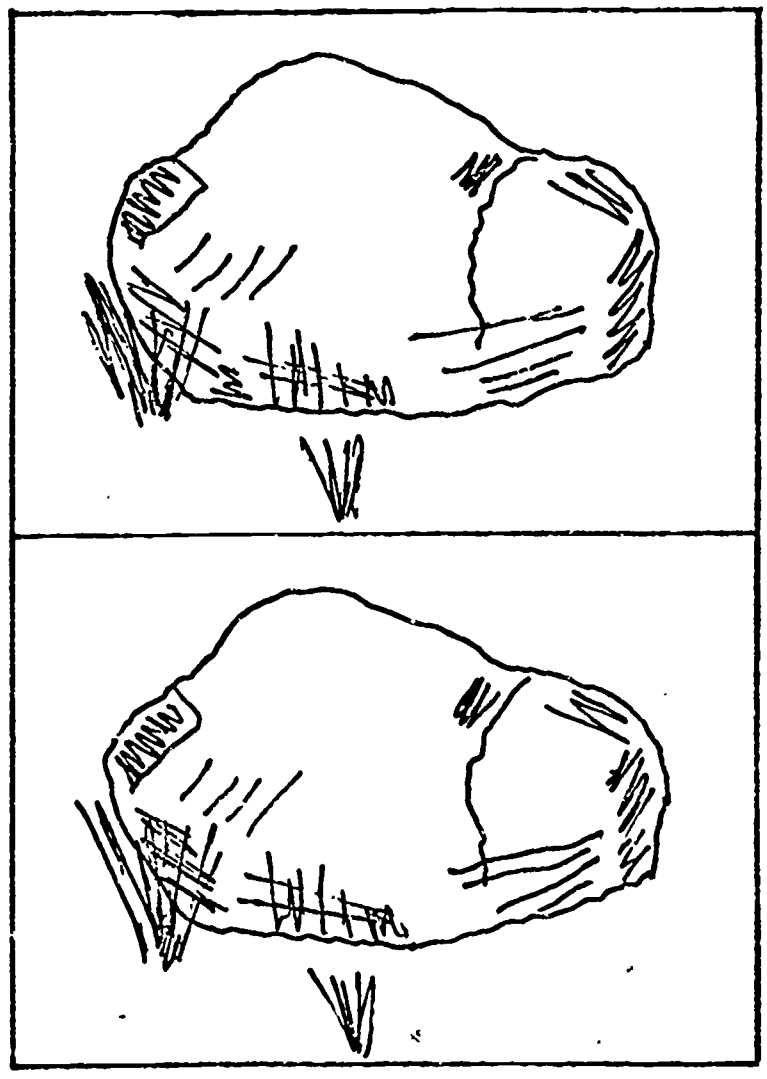




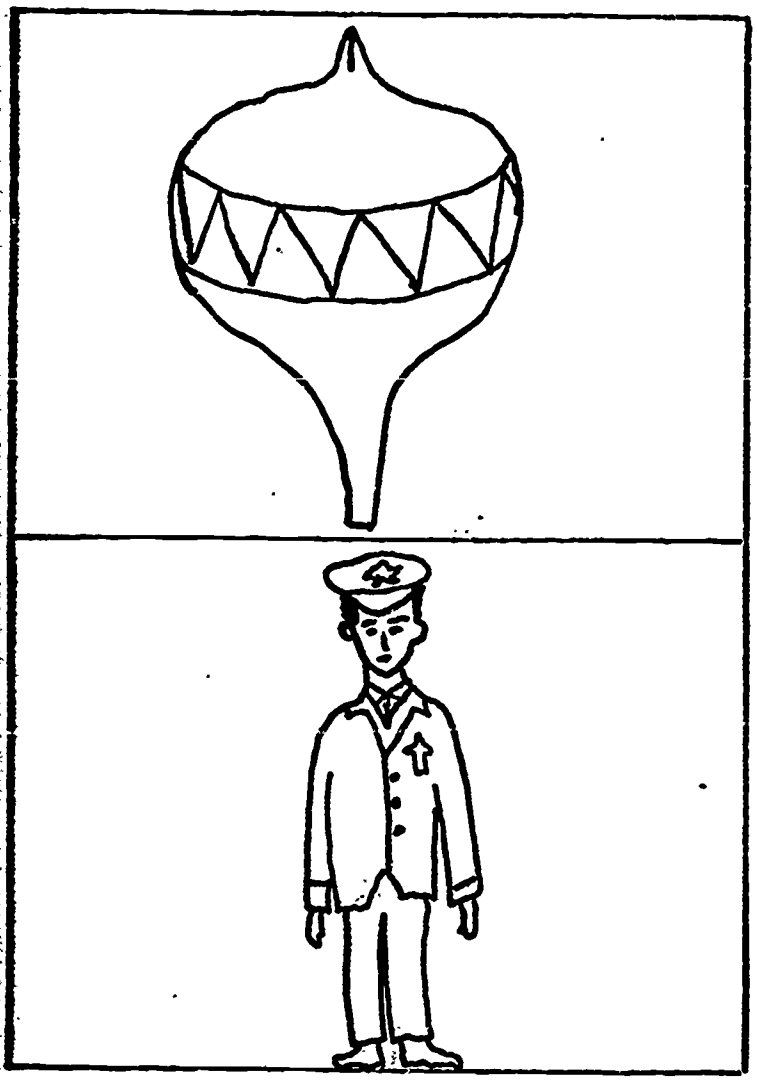


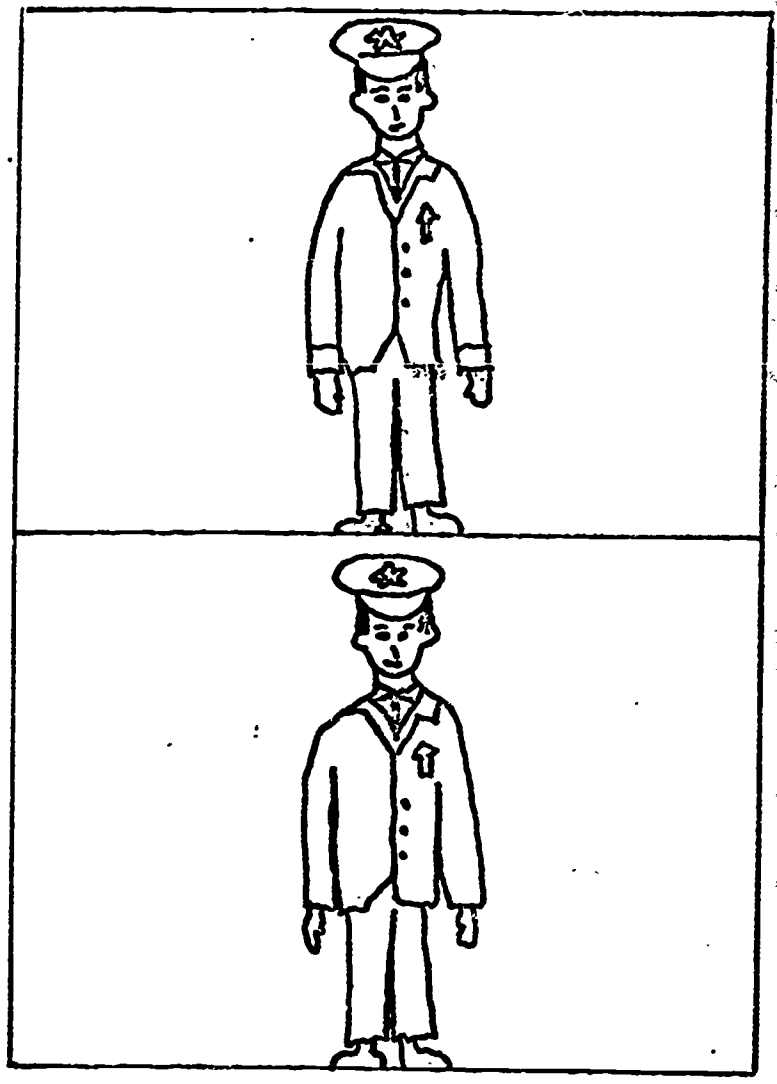




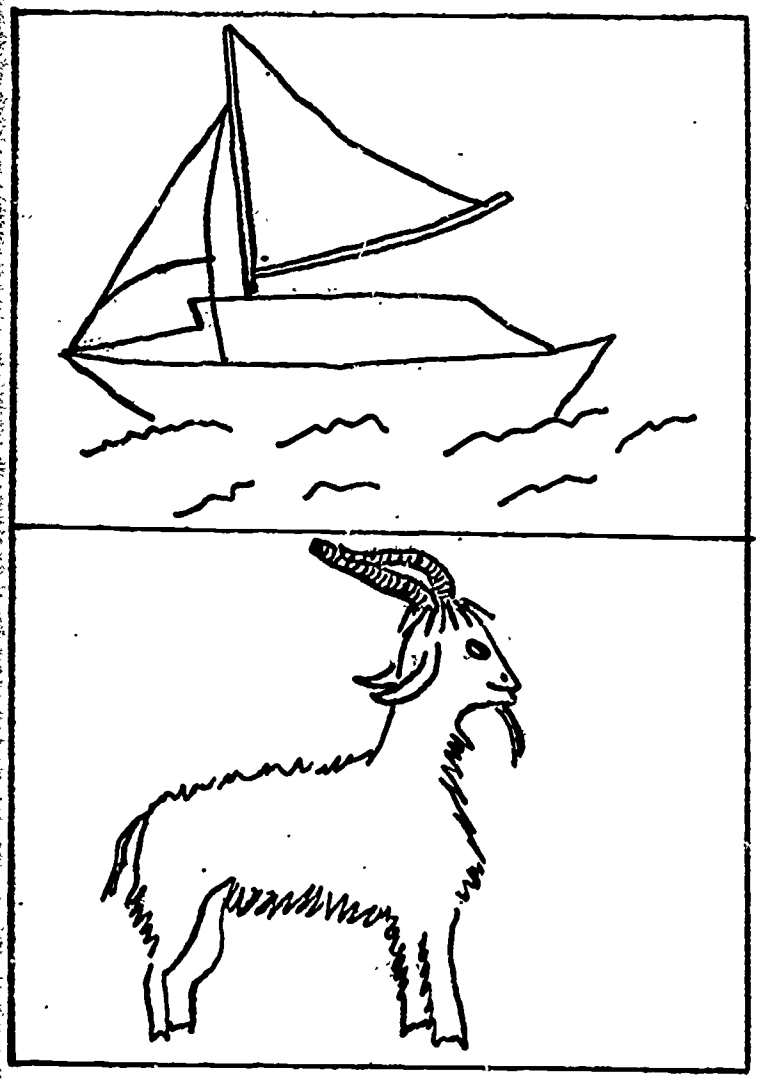


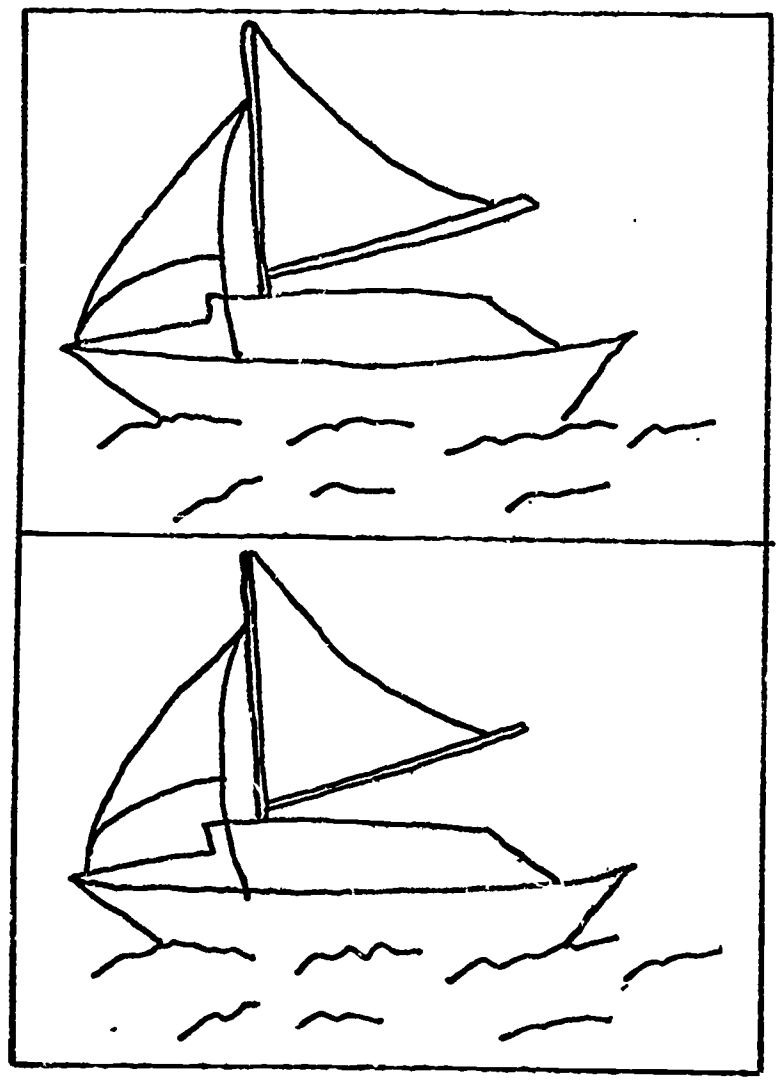
12





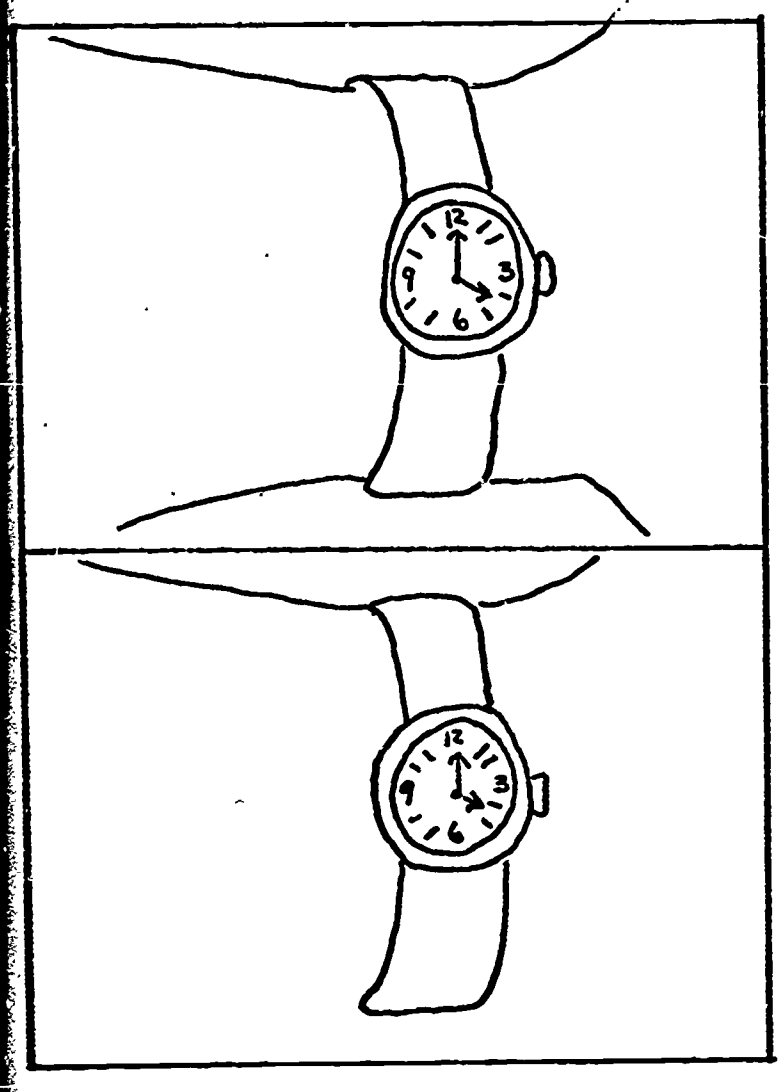
29

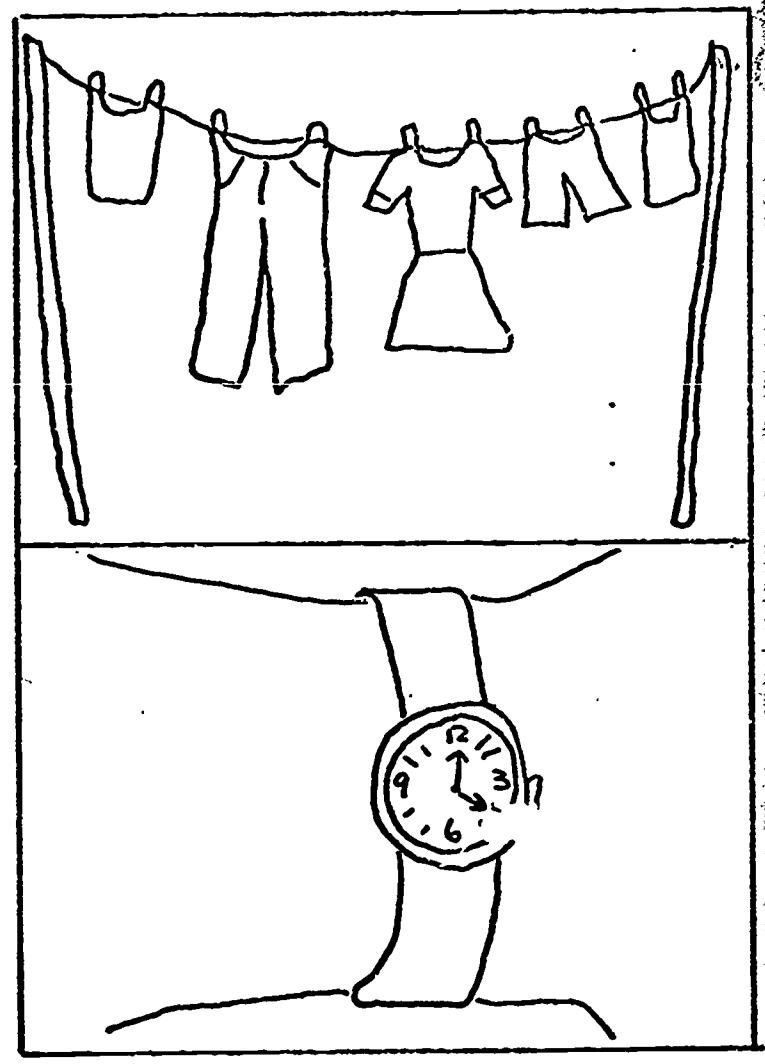




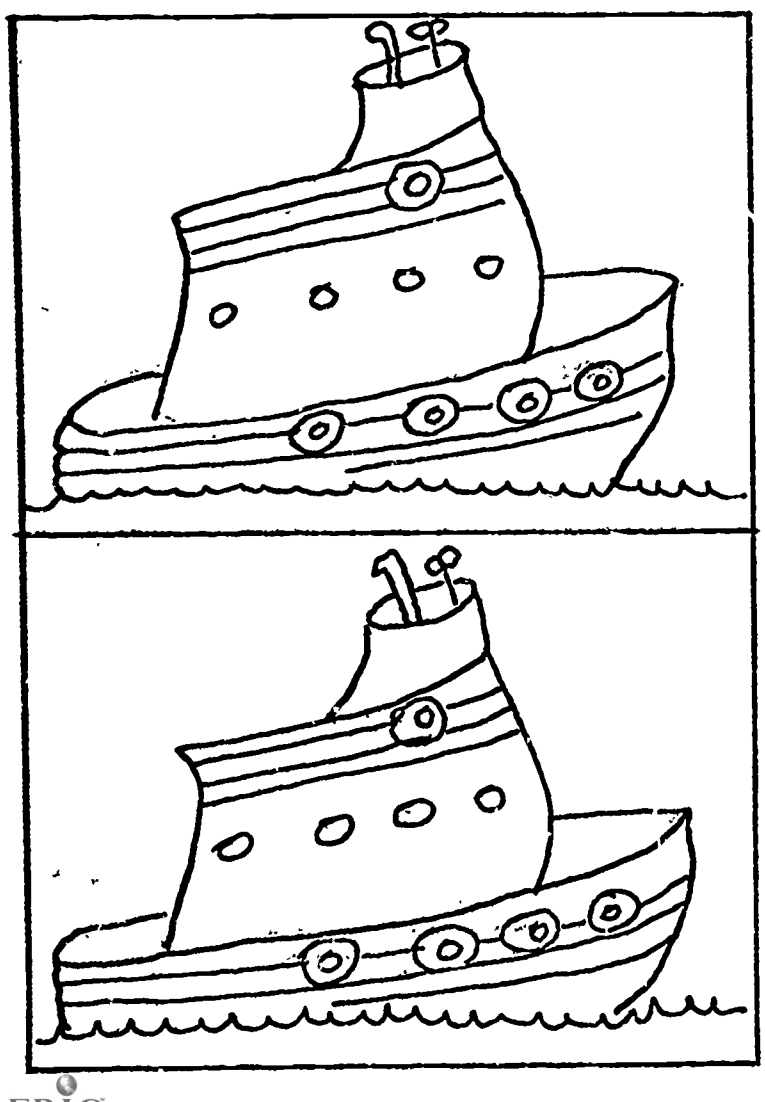
13

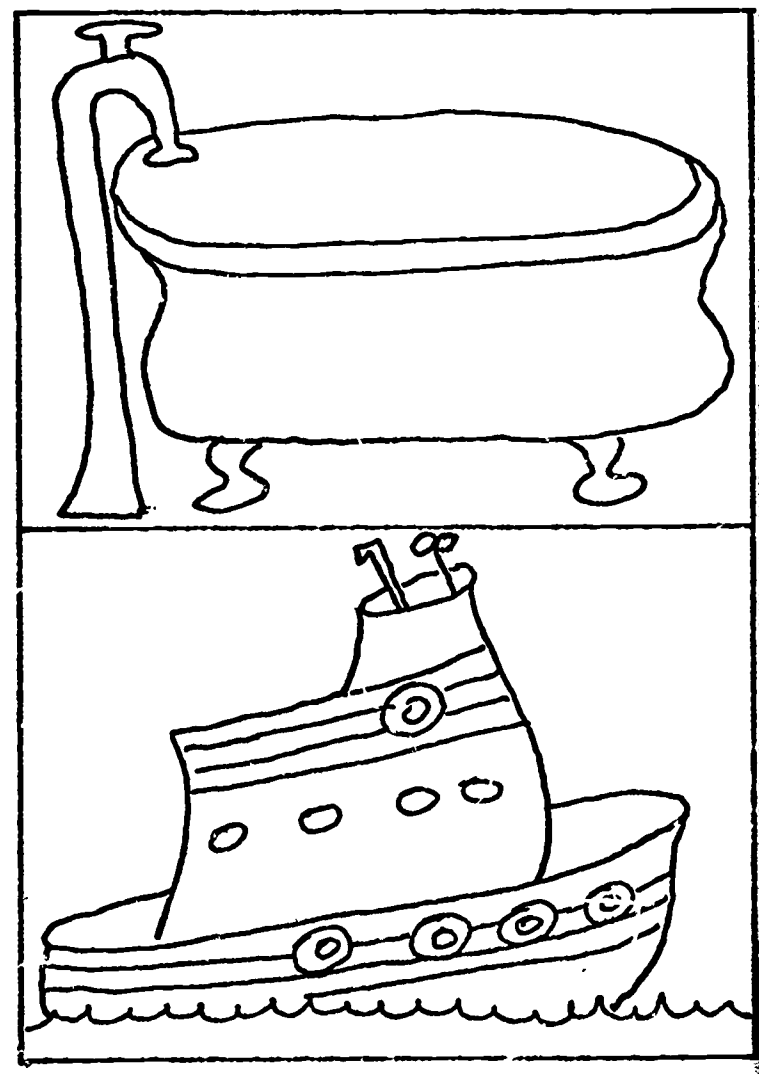
11



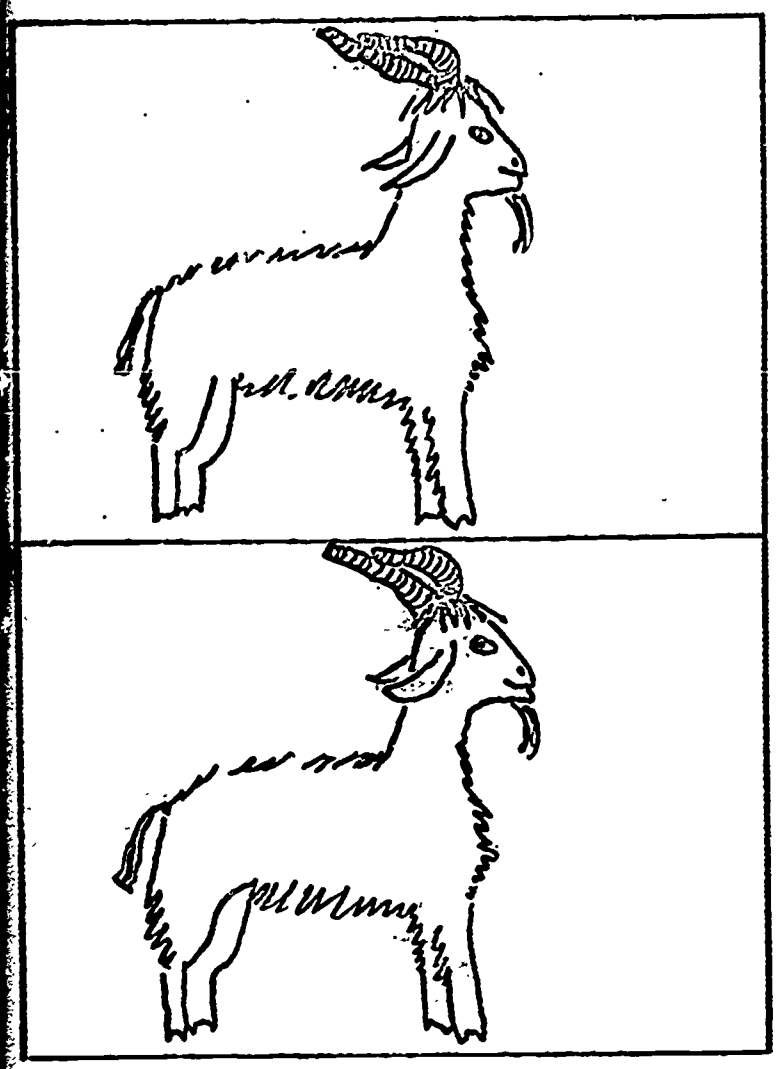


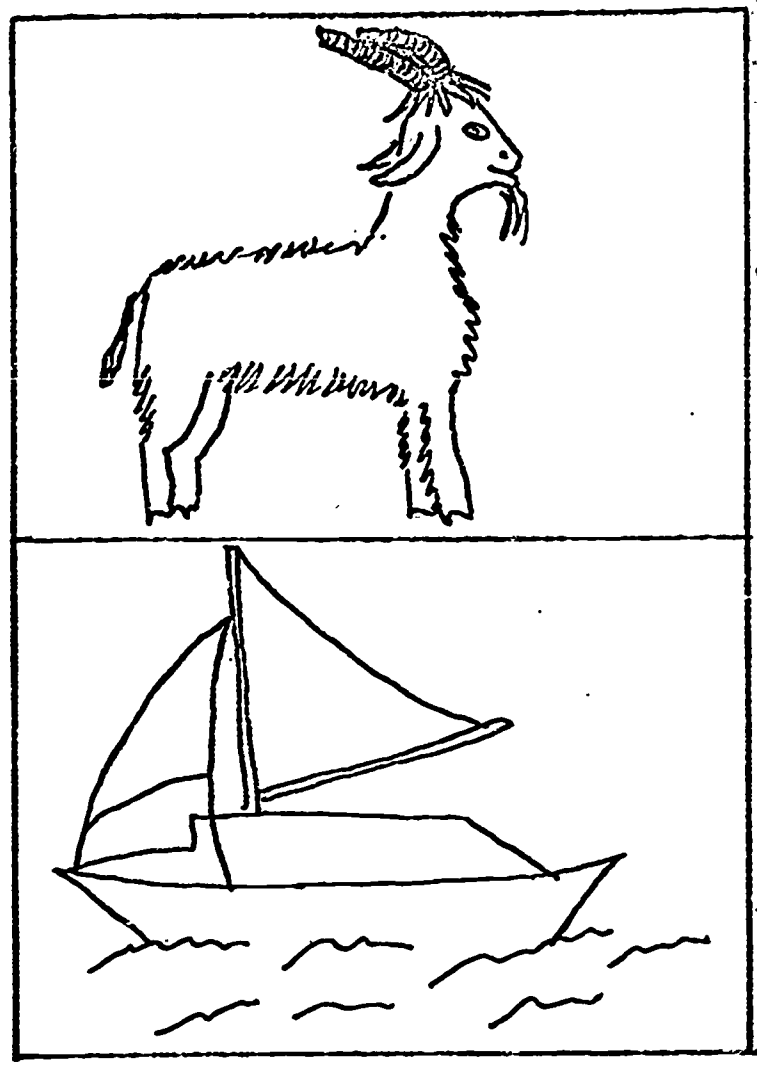
30



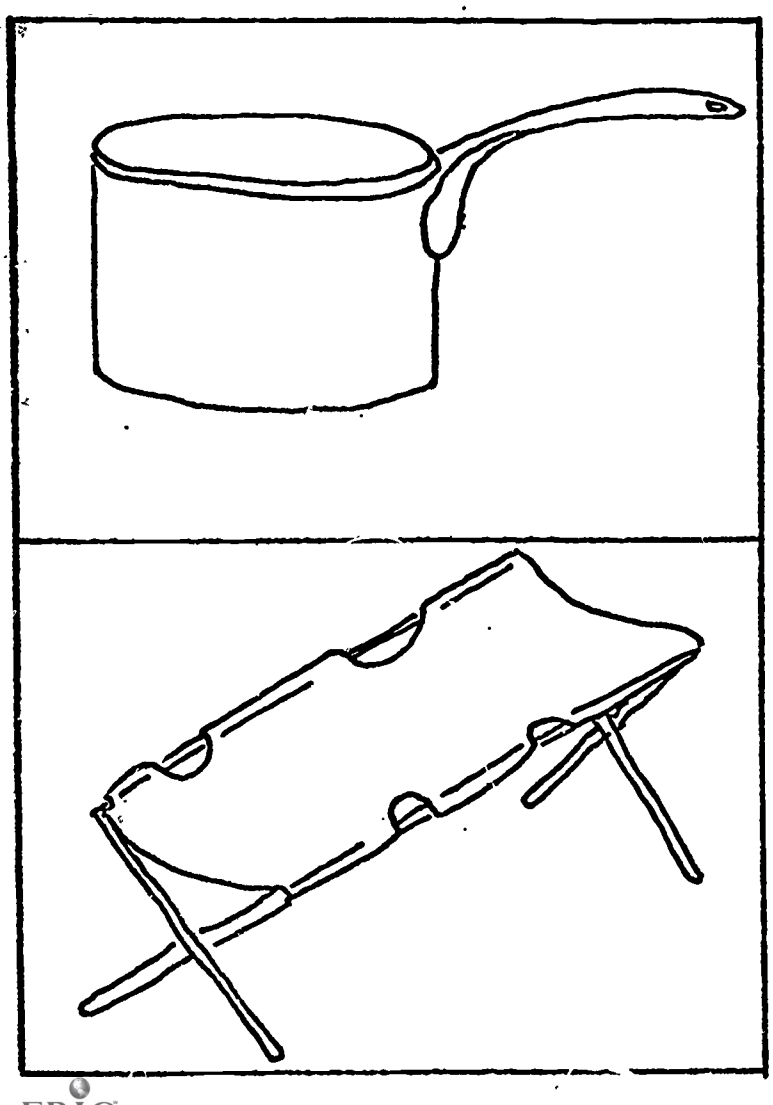


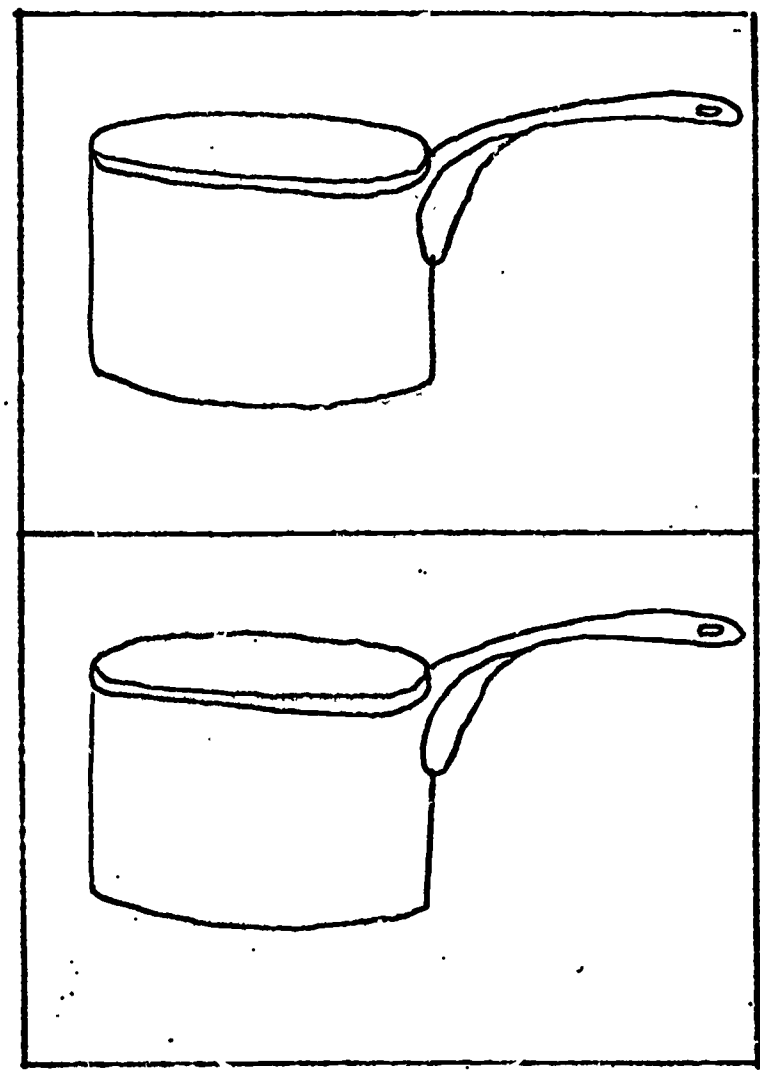
14

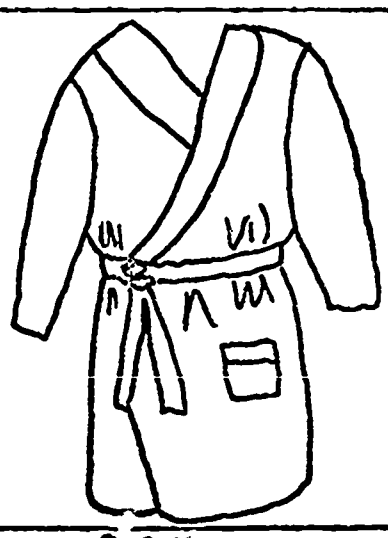


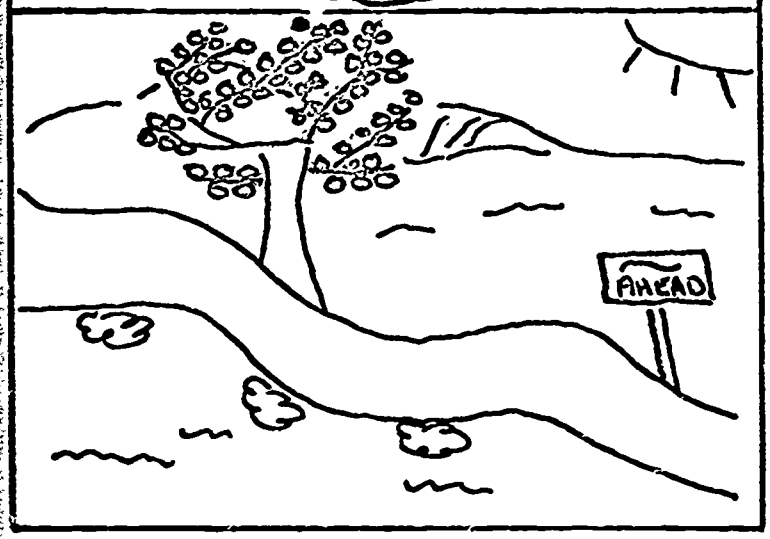


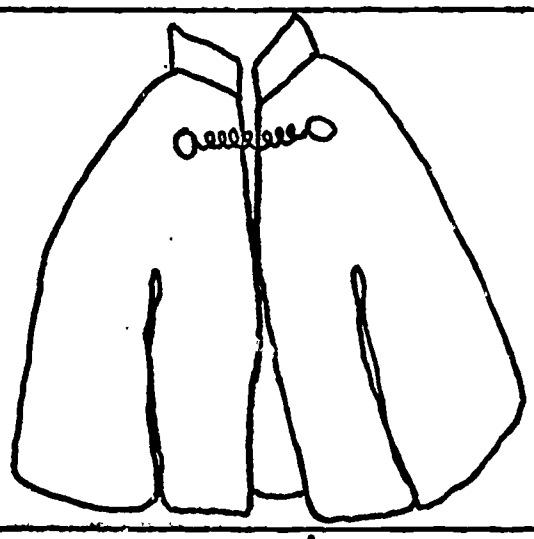
31

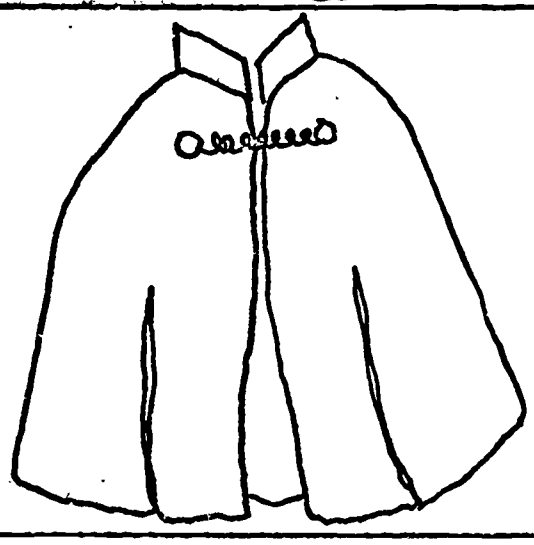


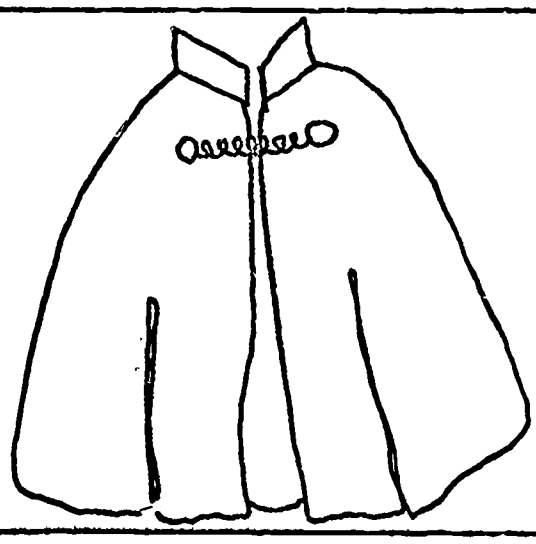


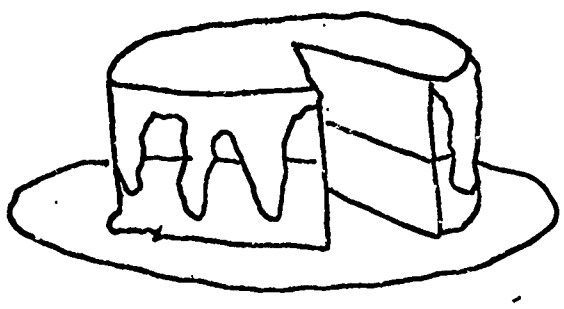



AHEAD

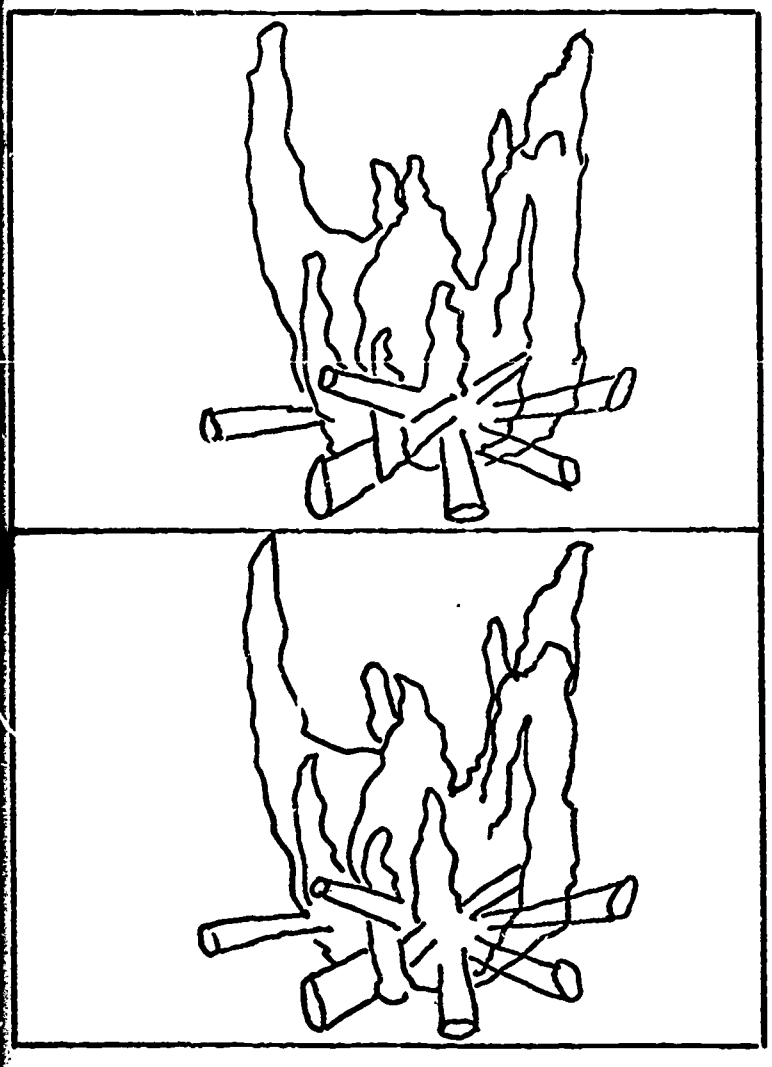


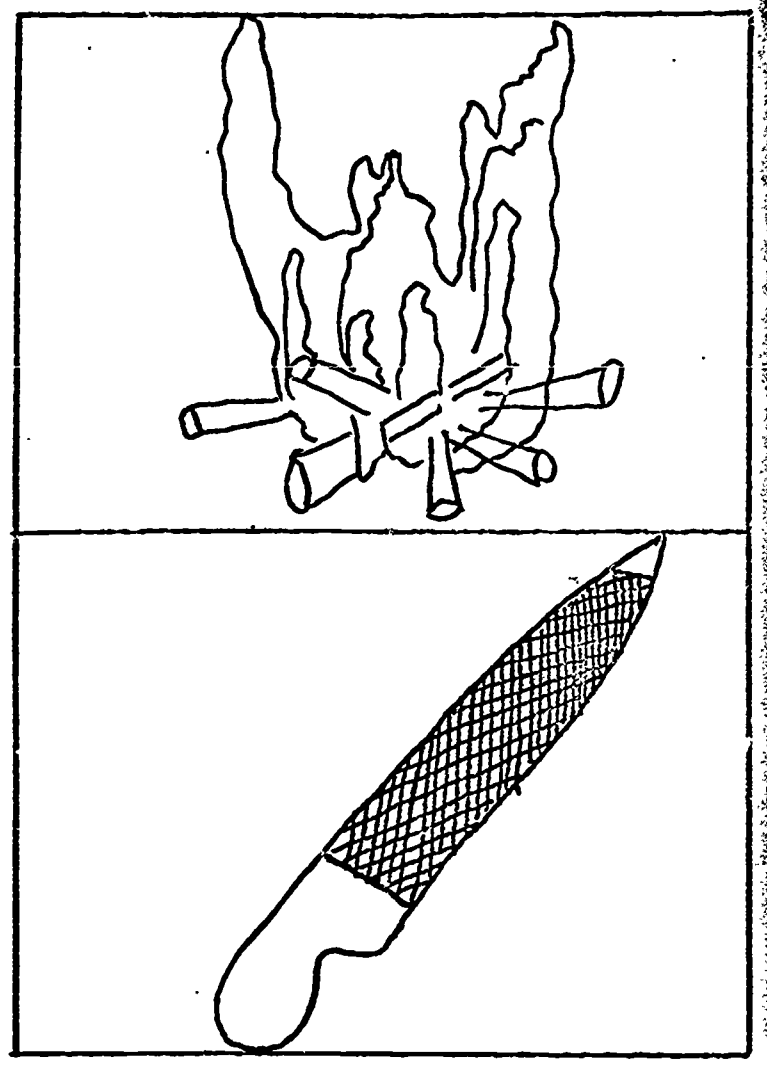






16





33