Journal of Comparative Psychology
1984, Vol. 98, No. 4, 381-404

A Vocabulary Test for Chimpanzees (*Pan troglodytes*)

R. Allen Gardner and Beatrix T. Gardner
University of Nevada, Reno

Chimpanzees can communicate in American Sign Language (ASL) to independent human observers whose only source of information is the ASL signs of the chimpanzees. A vocabulary test was presented to 4 cross-fostered chimpanzees (4-6 years old). Thirty-five-millimeter color slides were projected on a screen that could be seen by the chimpanzee subject but not by the human observers. There were two observers: O_1 was the questioner in the testing room with the subject; O_2 was in a different room. Neither observer could see the other, or the responses of the other observer. O_1 and O_2 agreed in their readings of both correct and incorrect signs, and most of the signs were the correct ASL names of the slides. In order to show that the chimpanzees were naming natural language categories—that the sign DOG could refer to any dog, FLOWER to any flower, SHOE to any shoe—each test trial was a first trial in that test slides were presented only once. Analysis of errors showed that two aspects of the signs, gestural form and conceptual category, governed the distribution of errors.

The critical role of early experience in the behavior of organisms is well-known and well-documented. Many animals have to learn to identify with their own species; many birds have to learn their mating songs. Even such habits as migration or wintering in the same place are profoundly influenced by species-typical rearing conditions (J. Brown, 1975, pp. 627-628, 643-652, 660). In cross-fostering, the young of one species are reared by foster parents of another species. So deep is the belief in the effect of rearing conditions on human behavior that even alleged but unverified cases of cross-fostering such as Itard's account of Victor, "the wild boy of Aveyron," attract serious scholarly attention.

In the twentieth century, Kellogg and Kellogg (1933) and Hayes and Hayes (1951) pioneered a form of cross-fostering in which the subjects were chimpanzees and the foster parents were human beings. In sharp contrast to their resemblance to young children in every other aspect of development, the chimpanzees acquired hardly any speech, even though the foster parents spoke to the adopted chimpanzees as parents speak to hearing children. The innovation of Project Washoe was to replace speech with American Sign Language (ASL), the manual language of the deaf in North America. Washoe learned ASL from her human companions and used signs in a childish and rudimentary way that resembled the early acquisition of speech and sign by human children (R. Brown, 1970; Hewes, 1973; Lieberman, 1984; Stokoe, 1978b; Van Cantfort & Rimpau, 1982). Of greater significance to the objectives of cross-fostering, two-way communication in ASL made Washoe's social and intellectual environment much more like that of human children.

Sign language in this laboratory was a means by which the chimpanzees could express their intelligence in ways that would permit comparison with human beings. The laboratory procedures assumed that all aspects of intellectual growth are intimately related. For young chimpanzees, no less than for human children, familiarity with simple tools such as keys, devices such as lights, articles of clothing such as shoes,

This research was supported by Grants GB 7432, GB 35586, BNS 75-17290, BNS 78-13058, and BNS 79-13832 from the National Science Foundation; Grants MH 12154 and MH 34953 from the National Institute of Mental Health; and grants from the National Geographic Society, the Grant Foundation, and the Spencer Foundation.

We wish to congratulate Benton J. Underwood on attaining his emeritus years and to express appreciation for his distinguished contributions to experimental psychology.

Requests for reprints should be sent to R. A. Gardner, Department of Psychology, University of Nevada, Reno, Nevada 89557.

are intimately involved in learning signs or words for keys, lights, shoes, opening, entering, lighting, and lacing. The laboratory was well stocked with such objects and activities, and the subjects had free access to them, or at least as much access as human children usually have. Although no more free than human children to go outdoors without permission, they were free of mechanical restraints both indoors and out. They not only learned to eat human-style food but also learned to use cups and spoons and to clear the table and wash the dishes after a meal. They not only learned to use human toilets (in their own quarters and elsewhere) but also learned to wipe themselves and flush the toilet, and even to ask to go to the potty in order to postpone lessons and bedtimes.

Our own observations (B. Gardner & Gardner, 1971, p. 141) and those of Hayes and Nissen (1971, written in 1957 and available to us in draft form in the early days of Project Washoe) taught us to avoid all forms of drill. In the Reno laboratory, the only time that formal, trial-by-trial procedures were used for teaching sign language (as opposed to testing) was in Fouts' (1972) experiment, a doctoral dissertation designed to isolate the effects of two specific procedures, modeling and molding.

Project Washoe was followed by a second project that included certain improvements. Washoe was captured wild and arrived in Reno when she was about 10 months old. Moja, Pili, Tatu, and Dar were born in American laboratories, and each arrived in Reno within a few days of birth. Their human companions had a high level of expertise in ASL. Many were deaf or the hearing offspring of deaf parents or had other extensive experience with ASL. Some were veterans of Project Washoe, and all had studied its procedures and results. Washoe, of course, was the only chimpanzee in her day, whereas the chimpanzees of the second project had each other as frequent companions.

Naturalistic Observation Versus Systematic Testing

Until recently, studies of the development of verbal behavior in human children concentrated on grammar to the near exclusion of vocabulary. The bulk of the evidence has been based on naturalistic methods such as diary records, inventories of phrases, and brief samples of utterances. The same naturalistic methods were used with Washoe, Moja, Pili, Tatu, and Dar, and comparisons can be found in B. Gardner and Gardner (1971, 1974, 1975, 1980; R. Gardner & Gardner, 1969; 1978). With respect to structure, Van Cantfort and Rimpau (1982) showed that when the same rules of evidence are applied to both sets of data, the same results are found for the early utterances of children and chimpanzees. Arguments for discontinuity that seem to be based on the same data have, instead, been based on a double standard of evidence. "Too often the comparisons that are cited are between three-year-old chimpanzees and university-level human beings or between observations of chimpanzee utterances and idealized, theoretical conceptions of human linguistic competence" (Van Cantfort & Rimpau, 1982, p. 65).

Aware of the limitations imposed by naturalistic observations, we supplemented our records with a program of systematic testing. Early versions of the vocabulary test that is the subject of this report are described in B. Gardner and Gardner (1971, pp. 158–161; 1974, pp. 11–15) and R. Gardner and Gardner (1974, pp. 160–161). The first objective of these tests was to demonstrate that the chimpanzee subjects could communicate information under conditions in which the only source of information available to a human observer was the signing of the chimpanzee. In order to accomplish this, pictures of nameable objects were projected on a screen that could be seen by the chimpanzee subject but could not be seen by the observer.

The second objective of these tests was to demonstrate that independent observers agreed with each other. In order to accomplish this, there were two observers. The first observer served as interlocutor in the testing room with the chimpanzee subject. The second observer, who was stationed in a second room, observed the subject from behind one-way glass. The two observers gave independent readings; they could not see each other, and they could not compare observations until after the test.

The third objective of these tests was to demonstrate that the chimpanzees used the signs to refer to natural language categories—that the sign DOG could refer to any dog, FLOWER to any flower, SHOE to any shoe, and so on (Rosch & Lloyd, 1978; Saltz, Dixon, Klein, & Becker, 1977; Saltz, Soller, & Sigel, 1972). This was accomplished by preparing a large library of slides to serve as exemplars. The first time that a particular chimpanzee subject saw any one of the test slides was on a test trial, and no test slide was shown on more than one test trial. Consequently, there was no way that a subject could get a correct score by memorizing particular pairs of exemplars and signs. That is to say, scores on these tests depended on the ability to name new exemplars of natural language categories.

Method

Apparatus and Procedure

Washoe

Apparatus. Figures 1A and 1B are diagrams of Washoe's testing apparatus. A slot in the wall (M.S.) between the testing room and the observation room permitted the first observer (O_1) to communicate in writing with the experimenter (E). The projection screen (P.S.) was so placed in the cabinet that when standing in proper position, O_1 could not see the image on the screen. The sliding door (S.D.) was locked closed except when O_1 was in the proper position and ready for the next trial. Only O_1 could unlock the sliding door with a special key. The sliding door had an additional latch that Washoe could operate, and it was counterweighted so that it opened automatically when Washoe opened her latch. When the door was open, a second observer (O_2) could view Washoe through a one-way glass window (1-w-G) built into the viewing cabinet. The window was so placed that O_2 could not see the projection screen A curtain (C) prevented O_2 from viewing E's data forms as well as any of the messages that had been passed to E either by O_1 or by O_2. A plastic barrier (P.B.) prevented Washoe from touching the screen or the window. The plastic barrier was so tilted as to polarize reflections from the projection screen that might be seen through O_2's window. A sheet of polarized plastic was placed over the window and so oriented as to prevent the polarized light reflected off the plastic barrier from passing through O_2's window.

Procedure. A trial started when Washoe opened the sliding door and thus exposed the picture on the projection screen. If Washoe failed to sign at this point, or if O_1 failed to read Washoe's sign for any reason, then O_1 could ask her to sign, sign again, or sign more clearly, depending on what seemed appropriate. If Washoe failed to make any scorable response, O_1 closed the sliding door to restart the trial. Usually, Washoe replied promptly with the name of an object. O_1 noted her reply, checked the screen in order to reward Washoe if she was correct, wrote down the English gloss for the sign on a slip of paper, passed the message in to E, and then closed the sliding door to prepare for the next trial.

In the observation room, O_2 also wrote down the English gloss for the sign that Washoe made, but O_2 never looked at the slide until after writing the gloss. The high level of agreement between O_1 and O_2, in their reports of incorrect as well as correct replies (Table 2 and Figure 7), was our safeguard against the possibility that looking at the screen before writing altered the reports of O_1. In addition, for Moja, Tatu, and Dar, O_1 and O_2 both wrote without looking at the screen with no noticeable loss in performance (Table 2 and Figures 7–9).

The messages were collected by E. If O_1 closed the sliding door without passing in a message, then E knew that the trial was not yet completed and showed the same slide again. When O_1 closed the sliding door and followed that with a message slip, the E knew that the trial was over and operated the carousel projector to show the next slide. A new trial began when Washoe reopened the sliding door to see the next picture.

Tatu and Dar

Apparatus. Figures 2A and 2B are photographs of Tatu's and Dar's testing apparatus. As in Washoe's apparatus, the projection screen was so placed within the cabinet that the image could not be seen from the O_1 position. A shutter on the projector replaced Washoe's sliding door, and a plug-and-jack switching device (S.P.) replaced the lock-and-key arrangement. When the plug was inserted into the jack, current was switched to a push button (B) that was set in the center of the threshold of the viewing cabinet. When so enabled, the subject could push the button to open the shutter and thus allow an image to appear on the projection screen. The shutter stayed open as long as the plug was inserted in the jack.

The one-way glass window that O_2 used was placed forward and beyond the plastic shield that protected the projection screen so that there was no surface that could reflect any image of the projection screen back to O_2. A wooden barrier prevented O_2 from seeing the projection screen through the one-way glass window.[1] The writing desk that E used was positioned so that the roof and walls of the viewing cabinet blocked O_2's view of E's data forms and earlier messages from O_1 and O_2.

---

[1] Motion picture films were made of Washoe and Dar under test conditions. When film was shot from the point of view of O_2, the one-way glass was removed. In Washoe's film the sheet of polarized plastic was also removed so that reflections sometimes show up on the plastic barrier.

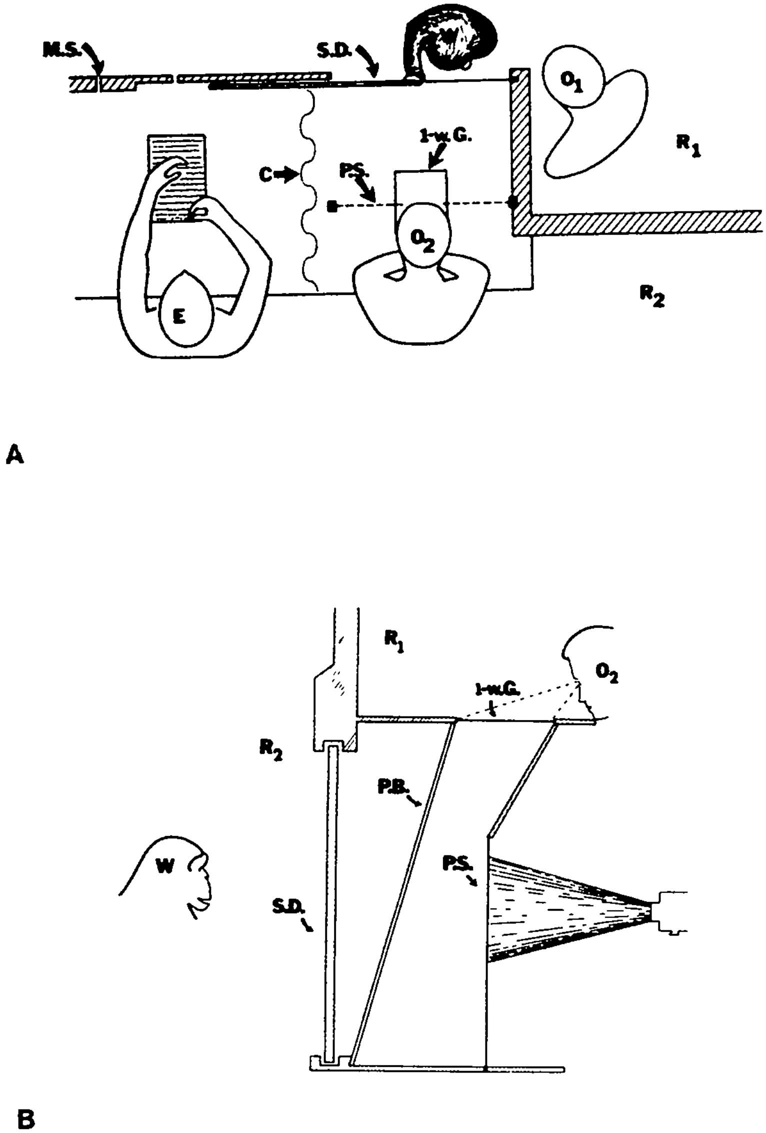

Figure 1 Vocabulary testing apparatus used in Project Washoe: Top view (A); side view (B). (M.S. = slot in wall between testing room [R_1] and observation room [R_2]; S.D. = sliding door; W = Washoe; O_1 and O_2 = observers; E = experimenter; C = curtain; P.S. = projection screen; 1-w-G. = one-way glass window; P.B = plastic barrier.)

Procedure. The procedure of presenting the pictures was the same as in Washoe's test except that O_1 locked and unlocked the shutter with the plug-and-jack switch and Tatu and Dar opened the shutter by pressing the button. As in Washoe's tests, O_1 wrote an English gloss for the chimpanzee signs on a slip of paper and passed it in to E through a slot in the wall between the testing room and the observation room. The procedure followed by E and by O_2 was the same as in Washoe's tests. Small treats were also given as rewards, but O_1 rewarded Tatu and Dar for signing promptly and clearly rather than for signing correctly. Thus, O_1 gave rewards immediately and always wrote the English glosses before looking at the projection screen.

Moja

The apparatus and procedure of Moja's tests were almost the same as those of Tatu's and Dar's tests except that the apparatus was less automatic. When Moja pushed the button, it only lighted a light as a

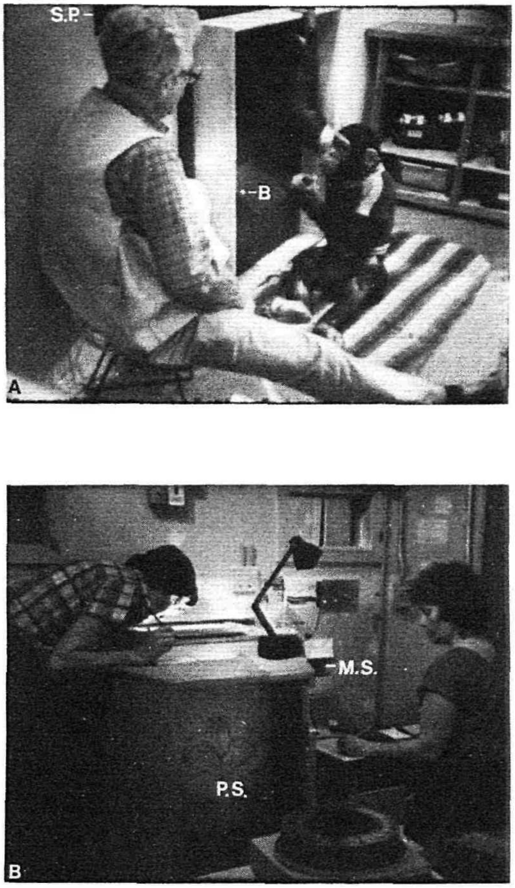

Figure 2. Vocabulary testing apparatus used with Tatu and Dar: Chimpanzee Dar and Observer 1 in Room 1 (A); Experimenter, right, and Observer 2, left, in Room 2 (B). (S.P. = switching device; B = push button; M.S. = slot between testing and observation rooms; P.S. = projection screen.)

signal to E to operate the shutter. As in the case of Tatu and Dar, O_1 gave Moja treats as rewards for signing promptly and clearly rather than for signing correctly. In Moja's case also, O_1 gave rewards immediately and always wrote the English glosses before looking at the projection screen.

Pretests

In most laboratory studies of nonhuman beings, the same procedures serve both for teaching and for testing. In the Wisconsin General Test Apparatus, for example, a monkey learns to associate one stimulus with the reward and the other with an empty food dish on the very same trials which are scored for correct and incorrect choices and plotted to show learning curves. Washoe, Moja, Tatu, and Dar were awake and accompanied by a human member of their foster family approximately 70 hr a week. During those 70 hr, the exposure to objects and the ASL names for objects was very great compared with the brief periods spent in vocabulary tests. Moreover, these tests were as different from the routines of the rest of their daily lives as similar testing would be for young children. For caged subjects, a session of testing is probably the most interesting thing that happens in the course of a laboratory day. For Washoe, Moja, Tatu, and Dar, most of the activities of daily life were more attractive than their formal tests. Nor could we starve them like rats or pigeons and make them earn their daily rations by taking tests.

Getting free-living, cross-fostered chimpanzees to do their best under the stringent conditions of these tests required a great deal of ingenuity and patience. A basic strategy was to establish the testing routine by a regular program of pretests that were kept short, usually less than 30 min, and infrequent, rarely more than two sessions per week. The purpose of the first series of pretests was (a) to pilot variations in procedure and (b) to ensure that experimenters, observers, and subjects were all highly practiced at the procedure before the first tests. The purpose of the second series of pretests was to ensure that they all stayed in practice while a new set of slides was prepared for the second tests.

Washoe

Washoe's earlier experience with blind testing is described in B. Gardner and Gardner (1971) and R. Gardner and Gardner (1973). Her first experience with slide tests began during the 38th month of the project when she was approximately 49 months old. After some pilot work, we designed and built the apparatus, and after several more adaptation sessions, the procedure described in the preceding section was put into practice in the 43rd month of Project Washoe.

From the 43rd to the 47th month, we conducted 19 pretests. Each pretest was completed in one session which lasted from 20 to 40 min. The pace for consecutive runs of trials was about two trials per minute, but O_1 kept the pace somewhat slower than Washoe seemed to want when she was doing well, and took time out for more active play when Washoe seemed to be doing poorly. Most of the pretests were 36–38 trials long and usually contained two examples of each of 16–19 different categories. About half of the slides on any given pretest were slides that Washoe had usually named correctly in previous pretests. The remaining slides were entirely new to Washoe or were slides that she had failed to name correctly on previous pretests and were being tried again. The order of presentation was unpredictable in that the experimenter who arranged the order shuffled the slides in a quasi-random fashion and neither O_1 nor O_2 knew the order in advance. Each order was presented forward on one pretest and then usually presented once more, only backward, on the next pretest.

During the 47th month of Project Washoe, we conducted the first test. From the 47th to the 50th month, we conducted 24 more pretests constructed in the same way as the first 19 pretests, and conducted Washoe's second test during the 51st month. Washoe left Reno at the end of the 51st month.

Tatu and Dar

After several adaptation sessions, the apparatus and procedure described in the previous section were used from Tatu's 54th month and Dar's 47th month. Dar's and Tatu's pretests were constructed in the same way as Washoe's except that (a) about one third of Tatu's pretests were less than 36 trials long, although the rest were 36–40 trials long, and most of Dar's pretests were 28–33 trials long and (b) the order of presentation was different for each pretest, the practice of presenting the same sequence forward and backward on two different pretests was rarely used, and for the last 27 pretests of both Tatu and Dar, the order of presentation was determined from a table of random numbers.

Tatu received 38 pretests from her 54th to her 59th month and her first test in her 59th month. She received 30 more pretests from her 60th to her 64th month and her second test in her 65th month. Tatu left Reno at the end of her 65th month. Dar received 41 pretests from his 47th to his 53rd month and his first test in his 53rd month. He received 28 more pretests from his 53rd to his 57th month and his second test in his 58th month Dar left Reno at the end of his 58th month.

Moja

The trials that Moja received before her test were equivalent to the adaptation sessions that the other subjects received before the formal procedures of their pretests were put into effect. Both Moja's performance and that of her observers must have suffered from lack of practice with the testing procedures.

Rewards

Early in Project Washoe, we learned that when Washoe was too anxious to earn her reward—when she was too hungry or the reward too desirable—then we could expect no more than the absolute minimum amount or quality of response necessary to get the reward. Whenever used, food rewards had to be very small—half of a raisin or a quarter of a peanut—more symbolic than nourishing. Attempting to reward Washoe for correct replies in the testing situation created procedural difficulties which we avoided with Moja, Tatu, and Dar by rewarding them for prompt, clear replies regardless of correctness. Unlike Washoe, the other subjects were distracted by the treats and would often ask for the rewards by name at critical points in the procedure, so that O_1 and O_2 could not tell whether the chimpanzees were asking for treats or naming a picture. Consequently, we abandoned this procedure entirely for Tatu and Dar, and they rarely received any edibles after the initial stages of pretesting.

Both the procedure in the observation room and the procedure in the testing room were frequently observed by additional personnel. Whether those serving as O_1 were aware of it or not, they often revealed their approval or disapproval by smiling or frowning and by nodding or shaking their heads as well as by signing such things as GOOD GIRL and SMART BOY.

Tests

Items and Exemplars

We use the term *vocabulary item* to refer to a category of objects, such as shoes, flowers, dogs, that could be named by a particular sign, and the term *exemplar* to refer to a unique member of such a category. All of the exemplars in the tests and pretests were 16 × 11 in. (41 × 31 cm) back-projected images of 35-mm color slides. When in earlier tests, exemplars were actual objects together with three-dimensional replicas (e.g., toy autos, ceramic birds) and photographs, Washoe named the photographs at least as well as she named the three-dimensional exemplars (B. Gardner & Gardner, 1971, pp. 158–161). Photographs are much more convenient, and this is particularly important for tests that must consist entirely of fresh exemplars.

Photography Outside the testing situation, we could use the pictures in ordinary books and magazines—even when the pictures contained objects of many different kinds—because we could point to particular objects or ask questions such as, WHAT BIRD EAT? Under the blind conditions of the tests, however, the objects had to dominate the field of view and the backgrounds had to be very plain. Even brightly colored backgrounds were sometimes distracting and got named instead of the objects. Note in the examples of Figures 3, 4, 5, and 6 that the objects usually filled the screen regardless of their normal relative size.

Some of the things the chimpanzees could name, such as grass or water, seemed impossible to present in test slides because they lack any characteristic shape and appear as meaningless forms against a blank background. Some otherwise shapeless objects, such as soda pop or facial tissues, could be presented in test slides because they appear in distinctive containers. Still other objects, such as houses and windows, seemed impossible to photograph without backgrounds or foregrounds that contained distracting extraneous detail. An important function of the pretests was to try out different photographic techniques. It was in this way that we learned to avoid vividly colored backgrounds that might be named instead of the featured object.

We had to learn to look at the slides with the eyes of our subjects. Thus, dramatic slides of leafy trees yielded mixed results, but shortly after Christmas one year, Tatu suggested something better. As she played with her discarded Christmas tree, she named it to herself many times. It turned out that Christmas trees, photographed close-up against the sky, made highly acceptable exemplars both for Tatu and for Dar. The tops of live evergreens taken with a telephoto lens were also highly acceptable. Even though evergreens come in some variety, we wanted to demonstrate that chimpanzees can also name deciduous trees under test conditions. Once more, Tatu showed us the way. That winter, on outings in the woods, Tatu frequently called our attention to trees by signing THAT TREE at their bare trunks. After this hint, we discovered that photographs of bare trees in winter made excellent exemplars.

Novelty. Each slide was unique and was shown only once on each test or pretest. An exemplar could

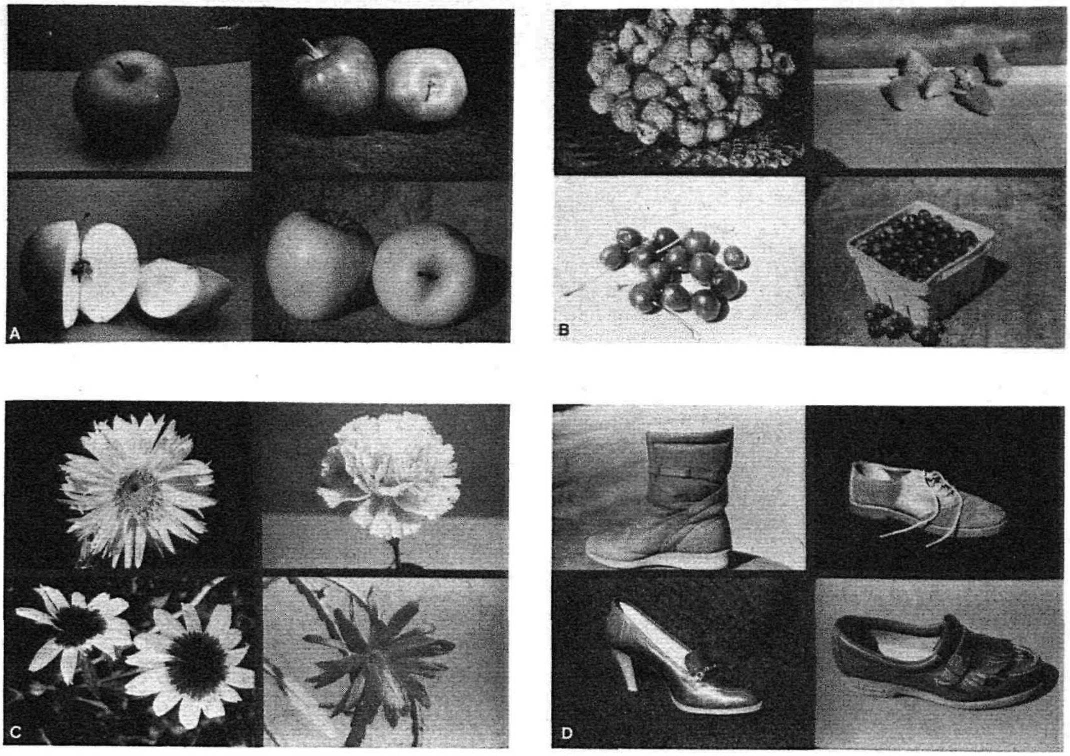

Figure 3. Black and white reproductions of exemplars of APPLE (A), BERRY (B), FLOWER (C), and SHOE (D) used in Tatu's Test 2 and Dar's Test 2.

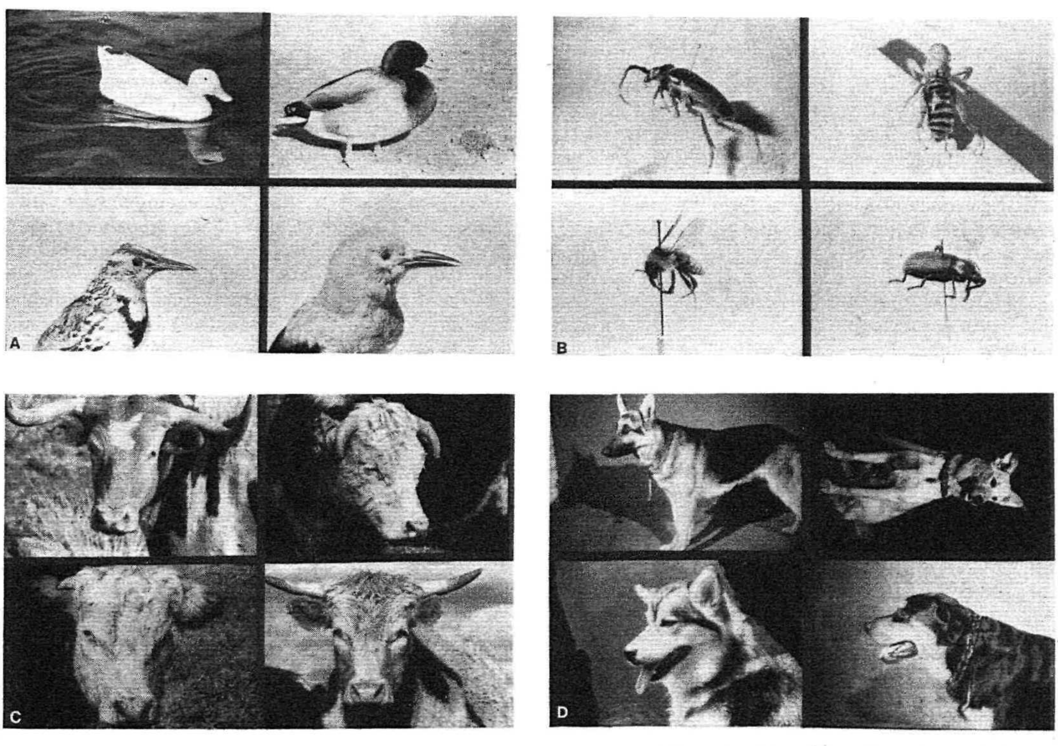

Figure 4. Black and white reproductions of exemplars of BIRD (A), BUG (B), COW (C), and DOG (D) used in Tatu's Test 2 and Dar's Test 2.

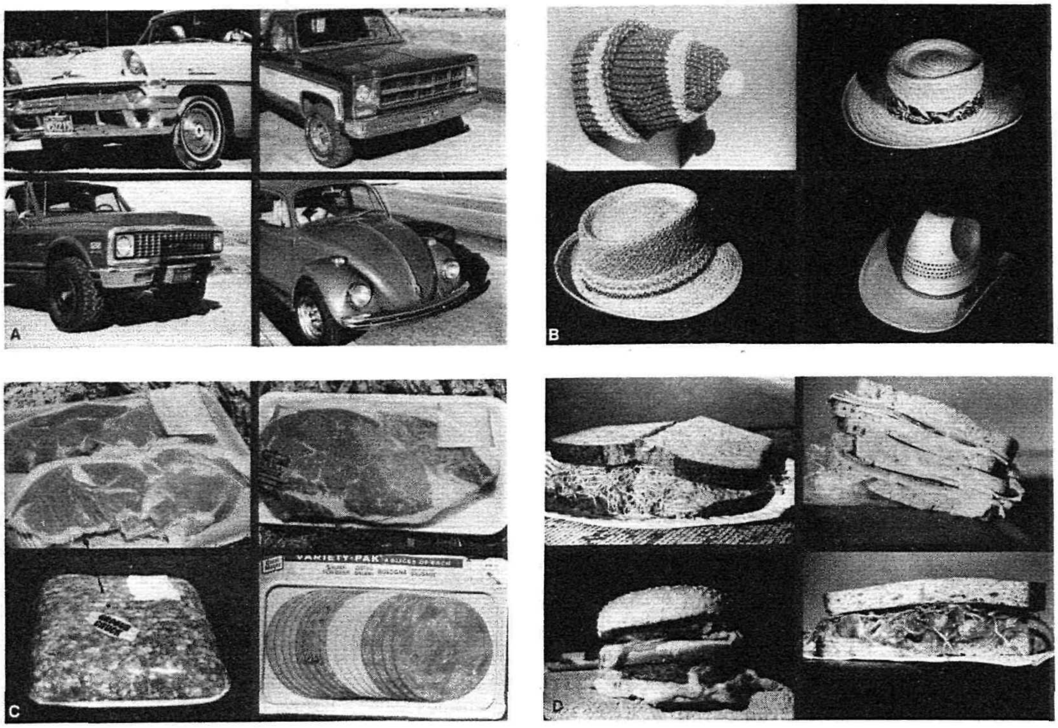

Figure 5. Black and white reproductions of exemplars of CAR (A), HAT (B), MEAT (C), and SANDWICH (D) used in Tatu's Test 2 and Dar's Test 2.

be unique because the chimpanzee subject had never seen that object before or because it was the first time that the subject had ever seen a photograph of that particular object or because it was the first time that the subject had ever seen that particular photograph of that particular object. When different exemplars consisted of different photographs of the same object, each was unique in that the object was photographed (a) from a different distance, (b) at a different angle, (c) against a different background, (d) under different light, or (e) in a different arrangement of a group, such as fruits, nuts, or shoes. At least three, but usually four, of the five dimensions were varied.

On Washoe's Test 1, 25% of the exemplars were photographs of objects that she had never seen before, either directly or as photographs, and 36% were new photographs of objects that she had never seen before except as photographs. On Washoe's Test 2, 40% of the exemplars were phtographs of objects that she had never seen before, either directly or as photographs, and 31% were new photographs of objects that she had never seen before except as photographs. On Moja's only test, 62% of the exemplars were photographs of objects that she had never seen before, either directly or as photographs, and 35% were objects that she had never seen before except as photographs. On the Test 1 of Tatu and Dar, all except three exemplars were photographs of objects that the chimpanzees had never seen before, either directly or in photographs, and those three exemplars were new photographs of objects that neither had seen before except as photographs. On Test 2 of Tatu and Dar, there was only one photograph of an object that they had ever seen before, either directly or in photographs. That one exception was a photograph of a longhorn steer that we included to add variety to the exemplars for COW. Neither Tatu nor Dar had ever seen any slide of that steer before, although they had seen the animal itself in a small herd in the neighborhood.

These requirements for uniqueness demanded a large library of slides. There had to be enough different slides to keep the pretests from becoming repetitious and to leave enough fresh slides to complete the two sets of tests with four widely differing exemplars of each vocabulary item. Moja's test slides were inferior to those of Washoe, Tatu, and Dar. They were prepared by an inexperienced photographer at a time when the laboratory was short-handed. We think that the inferior slides account for most of the difference between Moja's performance and that of the other subjects.

Composition

There were seven tests. Moja received only one test. The other 3 subjects each received two tests, and in each case the second test included the vocbulary items of the first test and additional vocabulary items. The

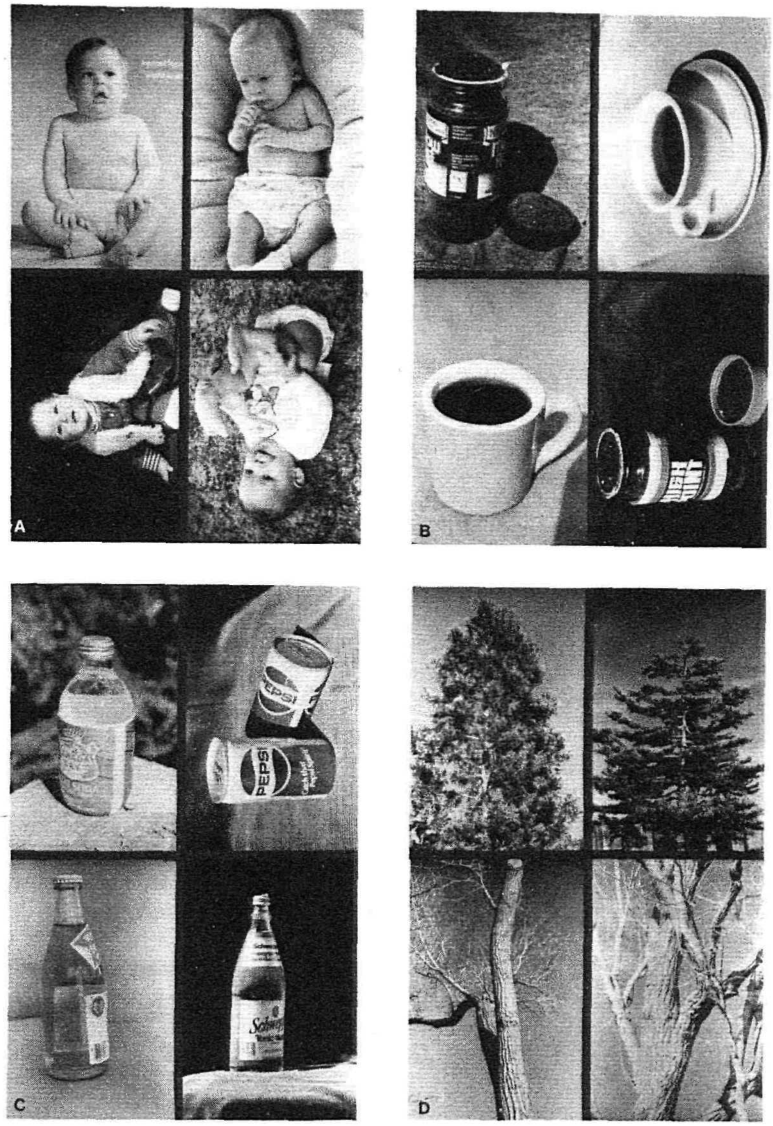

Figure 6. Black and white reproductions of exemplars of BABY (A), COFFEE (B), SODAPOP (C), and TREE (D) used in Tatu's Test 2 and Dar's Test 2.

vocabulary items that appeared in Moja's single test and in the second test of each of the other 3 subjects are listed in Table 1. Differences among the subjects, as shown in Table 1, reflect differences in their vocabularies, particularly differences between Washoe and the later subjects. Thus, Washoe had fewer signs for specific kinds of foods and drinks. Additional foods could have appeared in the tests of the later subjects, but we wished to avoid loading their tests with food items. We omitted some otherwise eligible vocabulary items, such as DIAPER and SWAB (really Q-tips) because the range of exemplars seemed too narrow.

The number of vocabulary items and the number of trials in each of the seven tests appear in Table 2. For each test, we chose four exemplars to illustrate the range of objects that a subject would name with the same sign. Different breeds represented CAT and DOG, different species represented BIRD and BUG, different makes and models represented CAR. The photographs that appear in Figures 3, 4, 5, and 6 illustrate the typical range of exemplars for vocabulary items.[2]

The sequence in which the exemplars were presented was determined from a table of random numbers without replacement and with the restriction that

[2] A complete set of the test slides can be obtained from the authors at current prices of duplication.

Table 1
Vocabulary Items in the Tests of Four Chimpanzees

Item	W	M	T	D	Item	W	M	T	D
Animates					Plants				
BABY	+		+	+	FLOWER	+	+	+	+
BIRD	+	+	+	+	LEAF	+	+		
BUG	+	+	+	+	TREE	+	+	+	+
CAT	+	+	+	+					
COW	+	+	+	+	Clothing				
DOG	+	+	+	+	CLOTHES	+			
HORSE		+			HAT	+	+	+	+
					PANTS	+			
Foods					SHOE	+	+	+	+
APPLE		+	+	+					
BANANA		+	+	+	Grooming				
BERRY		+	+	+	BRUSH	+	+	+	
CARROT			+	+	COMB		+	+	
CEREAL		+			HANKIE			+	
CHEESE	+		+	+	LIPSTICK			+	
CORN		+	+	+	OIL	+		+	+
FRUIT	+				TOOTHBRUSH	+	+	+	+
GRAPES		+			WIPER	+			
GUM		+							
ICE CREAM		+	+	+	Sensory				
MEAT	+		+	+	LISTENS	+	+		
NUT	+	+	+	+	LOOKS	+	+	+	
ONION		+			SMELLS	+			
ORANGE		+							
PEA-BEAN		+			Other				
PEACH		+	+	+	BALL			+	+
SANDWICH			+	+	BOOK	+	+		
TOMATO	+				CAR	+		+	+
					HAMMER	+			
Drinks					KEY	+	+	+	
COFFEE			+	+	KNIFE		+		
DRINK	+				PEEKABOO			+	
MILK		+			PIPE	+			
SODAPOP		+	+	+	SMOKE	+	+		

Note. W = Washoe; M = Moja; T = Tatu; D = Dar.

all four exemplars of the same vocabulary item could not appear in the same test session. Washoe's first test was divided into two equal test sessions. Each of the other tests was divided into four equal test sessions. Test sessions were separated by at least 2 days, and no more than two sessions were conducted during the same week. Immediately before the beginning of each test session, four easy trials were presented as criterion trials. The exemplars that were used on criterion trials were slides which that particular subject had always named correctly in pretests. If the subject named fewer than three of the four criterion exemplars correctly, the test session was aborted without showing any test exemplars. One session of Washoe's first test and one session of Tatu's first test were aborted for this reason. In addition, one session of Moja's test was ended after 19 trials because she failed to complete 20 trials within 35 minutes. The remainder of the trials scheduled for that session were presented in one short session 2 days later.

All observers were thoroughly familiar with the vocabularies of their subjects, but some of them were more involved than others in the construction of the tests and pretests. In order to reduce this source of variabilty, each observer was told in advance which vocabulary items could be represented by exemplars on each test and pretest. Tests, unlike pretests, required more than one session, and only a random sample of items appeared during any single session. Nevertheless, all observers reported additional signs that could not be scored as correct. That they reported extra-list intrusions is evidence that the observers were attempting to report the signs they saw rather than the signs that would give the chimpanzees the best possible scores.

Results

In order to fulfill the objectives of the tests and to calculate the scores expected by chance, a single correct target sign was designated for each picture in advance of the tests, and a single noun sign in each reply was designated as the response before the observers saw the picture on the screen.

Test Scores

Table 2 shows how the major objectives of the tests were accomplished. The agreement between O_1 and O_2 was high for all seven tests; except for Moja, the agreement

Table 2
Scores on the Vocabulary Tests of Four Chimpanzees

Measure	Washoe Test 1	Washoe Test 2	Moja Test 1	Tatu Test 1	Tatu Test 2	Dar Test 1	Dar Test 2
No. vocabulary items	16	32	35	25	34	21	27
No. trials	64	128	140	100	136	84	108
% interobserver agreement	95	86	70[a]	89	91	90	94
% scored correct							
By Observer 1	86	72	54[b]	84	80	79	83
By Observer 2	88	71	54[b]	85	79	80	81
Expected[c]	15	4	4	6	4	6	5

[a] Based on 135 trails; O_2 missed 5 trials. [b] Based on 132 trials; 8 unscorable trials. [c] Values based on the assumption that each observer was guessing with perfect memory.

ranged between 86% and 95%. Note that this is the agreement for both correct and incorrect trials. The agreement between the signs reported by the observers and the correct names of the categories is also high; except for Moja, correct scores ranged between 71% and 88%.

Targets

The correct sign for each vocabulary item was designated in advance of the tests. That sign and that sign only was scored as correct for that item, with two exceptions. The sign that is usually glossed as TOMATO in ASL is a compound that can be literally translated as RED SLICES (Stokoe, Casterline, & Croneberg, 1976, p. 157). In informal conversation, native signers commonly sign either RED or SLICES alone as equivalents for this compound. Outside the testing situation, Washoe sometimes called tomatoes RED and sometimes RED SLICES, and we accepted both as equivalents for TOMATO in Washoe's tests and pretests. The sign that is usually glossed as SODAPOP is also a compound (Stokoe et al., 1976, p. 216), and native signers commonly use the second element of the compound, literally POP, as the equivalent of SODAPOP. Outside the testing situation, Moja, Tatu, and Dar signed POP and SODAPOP as equivalents, and we accepted both as equivalents for SODAPOP on their tests and pretests.

Replies

Most of the replies consisted of a single noun sign naming an object, and that was the sign that was scored as correct or incorrect with respect to the target. On some trials the single noun in the reply was contained in a descriptive phrase, as when Tatu signed RED BERRY for a picture of cherries or when Dar signed THAT BIRD for a pintail duck. Because these replies contained only one noun sign, that was the sign that was scored as correct or incorrect with respect to the target. Although there are aspects of the pictures for which superordinate terms, such as FOOD, or descriptive terms, such as BLACK, might be scored, neither the presence or absence nor the correct or incorrect use of such terms was considered in scoring these tests.

In ordinary conversation, replies that consisted entirely of superordinate or descriptive terms and phrases were often appropriate and acceptable. Because such replies could not be scored as correct on test trials, O_1 attempted to induce the subject to be more specific by asking a probing question, such as WHAT KIND FOOD? or WHAT BLACK?, or O_1 might restart the trial. Because overall performance on the tests depended on the good humor of the subjects, there were severe limits to the amount of such probing and restarting of trials.

On 14% of the test trials of Washoe, Tatu, and Dar, either O_1 or O_2 reported more than one object name. On these trials the observers had to choose one of the names for scoring before seeing the picture themselves. The agreement between O_1 and O_2 on these trials was 69%, and the chimpanzees were correct on 61% of O_1's reports and 68% of O_2's reports. This is well below the overall performance of these subjects that is reported in Table 2. Initially, the sign to be scored on such trials was designated as the first object name in the reply. This was the rule followed outside of the testing situation to discourage indiscriminate guessing. Hoping to raise test scores, O_1 and O_2 were given the option of designating an object name other than the first, if it was signed more emphatically or more confidently. Again excluding Moja, either O_1 or O_2 used this option on 8.7% of all test trials. Considering those trials alone, interobserver agreement was 72% correct; correct score by O_1's report was 78%, and correct by O_2's report 79%, but these results are still poorer than the overall performance reported in Table 2.

Thus, on those trials in which the chimpanzees replied with more than one object name, communication was impaired and scores were lower. Perhaps the subjects were indecisive because they were confused by some of the pictures; perhaps they were relatively inattentive on some trials. As expected, Moja's observers reported more indecisive trials, more disagreement between observers, and lower scores.

Chance Expectancy

In B. Gardner and Gardner (1971), we estimated the expected chance score as 1/

N where N is the number of vocabulary items on a test and all items are represented by the same number of exemplars. This estimate assumes that only the chimpanzees were guessing. It may be too low because it does not take into account the possibility that the observers were guessing. Both O_1 and O_2 were familiar with the list of items that could be represented in a test and often knew which slides had been shown on past trials. In random sampling without replacement, the probabilities of later events in a sequence depend on earlier events. The observers could have used their knowledge of the items that had appeared earlier to predict the items that would appear later. Thus, players who can remember the cards that have already been played can win significant amounts at games such as Black Jack. Diaconis (1978) and Read (1962) deal with a similar problem in demonstrations of extrasensory perception (ESP). When highly motivated subjects in ESP experiments can see each target card after each prediction, their later predictions tend to improve.

To estimate the effect of informed guessing by the observers on chance expectancy in the tests reported here, Patterson (1983) performed a computer simulation which assumes that both observers (a) saw each slide after each trial, (b) had perfect memory for the number of exemplars of each vocabulary item that had appeared before the beginning of each trial, and (c) guessed the correct sign on the basis of the number of exemplars of each vocabulary item that remained to be presented. The line labeled "Expected" in Table 2 shows the average results of 2,000 simulated runs for each of the seven tests. In all cases, this estimate is a small fraction of the obtained scores. Since, O_1 and O_2 reported extralist intrusions (signs that were not on the target lists), they were using a less efficient strategy than that assumed in Patterson's simulation. Hence, low as they are, the values in the "Expected" line of Table 2 overestimate chance expectancy.[3]

The expected score for Washoe's first test is appreciably higher than the expected scores for the other six tests for two reasons. First, this test was shorter than the other tests, and predictability depends on the number of vocabulary items—the fewer the items the greater the predictability. Second, and more significant for this discussion, predictability increases as one approaches the end of the test. The last trial is completely predictable, because there is only one vocabulary item that could have any remaining exemplars. The next to the last trial may be completely predictable, but there are at most two vocabulary items that could still appear, and so on. In all cases except for Washoe's first test, both O_1 and O_2 were assigned to test sessions in such a way that no individual served as an observer for more than half of the trials of any single test, and of course, no observer saw any trials other than those assigned until after serving in the assigned test sessions. The device is similar to the way gambling casinos can defeat card-counting customers by frequently reshuffling the deck. The smaller number of items and the assignment of the same two observers to both sessions of Washoe's first test account for the higher, but still quite small, expected value on that test.

The suggestion that the results of these tests could be attributed to such a negligibly small opportunity for the observers to guess the correct signs (Umiker-Sebeok & Sebeok, 1980, p. 40) fails to consider the enormous difference from gambling or extrasensory perception. In order to win at a game such as Black Jack, the card-counter does not have to guess the exact values of the remaining cards, the way the observers on these tests had to match exact vocabulary items. All the card-counters have to do to win money from the house is to guess whether more high or more low cards remain in the deck and then raise their bets when the odds are in their favor or lower their bets when the odds are unfavorable. For this purpose, quite small deviations from chance expectancy can be parlayed into significant winnings. In the case of extrasensory perception, the objective of

[3] Requests for copies of this program and results of the computer simulation and for copies of the proof of the generalization of Read's formula for the expected number of correct guesses can be sent to J. C. Patterson, Department of Mathematics, Eastern Kentucky University, Richmond, Kentucky 40475.

the experiment is to demonstrate the existence of powers that would be truly discontinuous with the rest of psychology and biology. For the purpose of demonstrating such a discontinuity, any statistically significant deviations from chance—however small—would have profound implications.

The ability of chimpanzees to grasp concepts such as those listed in Table 1 and to communicate that information by means of gestures is continuous with the rest of comparative psychology. The evidence consists of large departures from chance expectancy, such as the scores in Table 2. The effects are of a much greater order of magnitude than those that might be considered significant in gambling or demonstrations of ESP.

Signs of ASL

Independent observers agreed in their reports of chimpanzee signs, and the signs that they reported were the correct names for the objects in the pictures. Written descriptions of these signs were part of the daily laboratory procedure. (R. Gardner & Gardner, 1978, pp. 46–49; B. Gardner & Gardner, 1980, pp. 340–341, 350–351). These descriptions were frequently and thoroughly checked and rechecked and compared with published descriptions (e.g., Fant, 1977; Stokoe et al., 1976; Watson, 1973). The agreement with ASL was additionally verified by consultation with outside experts who reviewed the written descriptions and photographic examples and who also observed the chimpanzees in person (e.g., Stokoe, 1983).

We used the test procedures for further verification. Two fluent deaf signers, both then recent graduates of Gallaudet College, each served as O_2 in Washoe's pretests in the summer of 1970. Each of these young men participated in two pretests at a time when each had observed Washoe for less than 1 hr. Their agreement with O_1 (who had in each case years of experience with Washoe) rose from 67% and 71% on their first session to 89% for both on their second session. The task that these fluent signers faced should be compared with that of fluent speakers of English identifying the words of equally immature human children after equally brief preexposure to the immature speakers and under equally stringent conditions. These two outside observers who were expert at ASL, but unfamiliar with Washoe, read her signs fairly well at first and then improved markedly in their second test session. The improvement did not depend on learning about specific vocabulary items because (a) items were changed from the first session to the second session and (b) the deaf observers were told in advance which items to expect, just as the regular observers were. The initially good agreement with O_1, together with the improvement, indicates that Washoe's signs were intelligible to fluent signers with, perhaps, a childish or chimpanzee accent that could be learned fairly quickly.

Distribution of Errors

With respect to the agreement of the observers with each other and the agreement of their reports with the correct names of the exemplars, all of the trials can be assigned to one of four categories: (a) agreed correct, in which the same sign was reported by both observers and that sign was correct; (b) agreed errors, in which the same sign was reported by both observers and that sign was an error; (c) half errors, in which one observer reported the correct sign, but the other observer reported a different, necessarily incorrect sign; and (d) disagreed errors, in which O_1 and O_2 each reported a different incorrect sign.

Because interobserver agreement was high (Table 2) and because agreement is itself evidence of reliability, we confined this analysis to the category of agreed errors. Because this analysis involves contingency tables with many cells and because there were so few errors reported during Tests 1 and 2, we pooled the agreed errors of Tests 1 and 2 with the agreed errors of the pretests to obtain the data that appear in Figures 7, 8, and 9. The rows of Figures 7, 8, and 9 represent the signs reported as agreed errors during the pretests and Tests 1 and 2 for Washoe, Tatu, and Dar, respectively; the columns represent the targets, that is, the names that we assigned to the slides that were presented on those trials.

	Targets	IL				OL					MF			HD					HN						PE								
Errors		DRINK	NUT	PIPE	TOOTHBRUSH	TOMATO	SMOKE	BIRD	WIPER	SMELL	FLOWER	BUG	LOOK	COW	HAT	CAT	FRUIT	LISTEN	KEY	CHEESE	BOOK	MEAT	OIL	SHOE	BRUSH	BABY	LEAF	TREE	CLOTHES	PANTS	DOG	HAMMER	CAR
IL	DRINK	.	(1)	15	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	NUT	(1)	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	PIPE	23	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	TOOTHBRUSH	2	.	.	.	.	1	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	FOOD	(1)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OL	TOMATO	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	(1)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	WATER	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	SMOKE	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	BIRD	.	.	.	.	2	.	.	1	.	.	.	.	.	.	(9)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	LOLLY	(1)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	WIPER	.	.	.	.	.	1	.	.	7	.	.	.	.	1	.	5	1	.	1	.	.	.	.	(1)	.	.	.	.	.	.	.	.
	SMELL	.	.	.	.	.	.	.	5	.	?	1	.	.	1	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.
MF	FLOWER	.	.	.	.	.	.	.	.	2	.	20	1	.	.	.	1	.	.	.	1	.	.	.	.	.	.	.	.	.	.	1	.
	BUG	.	.	.	.	.	.	.	.	1	6	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	LOOK	.	.	.	.	.	.	1	.	.	3	.	.	.	1	.	1	(2)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HD	MAN	.	.	.	.	.	.	.	.	.	.	.	.	.	.	(1)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	COW	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	HAT	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.
	COMB	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	(1)	.	.	.	.	.	.	.	.
	CAT	.	.	.	.	.	.	(11)	.	.	.	.	.	.	.	.	2	.	.	.	.	.	.	.	.	.	.	.	.	.	(4)	1	.
	FRUIT	.	.	.	.	(3)	.	.	2	1	1	.	1	.	3	2	.	.	.	(3)	1	.	.	1	2	.	.	1	.	.	2	.	1
	BED	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	LISTEN	.	.	.	.	.	.	.	.	.	1	.	(5)	.	.	.	.	.	.	.	.	.	.	2	.	.	.	.	.	.	.	1	.
	DIRTY	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.
HN	PENCIL	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	2	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	SPOON	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	2	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	KEY	.	.	.	.	.	.	.	.	.	.	1	3	.	.	.	.	1	.	.	.	1	.	.	.	.	1	.	.	.	.	.	.
	CLEAN	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	CHEESE	(1)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	(1)	.	.	.	.	(1)	.	.	.	.	.	.	.	.	.	.	.
	BOOK	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	2	.	.	.	.	.	.	.	.	2	.
	MEAT	(1)	.	.	.	.	.	1	.	.	.	.	.	.	.	.	(1)	.	.	.	.	.	2	.	.	.	.	.	.	.	.	.	.
	ROPE	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.
	OIL	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	7	.	.	.	.	.	.	.	.	.	.	.
	WINDOW	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	3	.	.	.	1	.	.	.	.	.
	SHOE	.	.	.	.	.	.	.	.	.	.	.	.	.	(1)	.	.	.	.	.	.	.	.	.	.	.	.	.	(1)	.	2	1	.
PE	BRUSH	.	.	.	.	.	.	.	(2)	.	.	.	.	.	3	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	2	.	.
	BLANKET	.	.	.	.	.	.	.	.	.	.	.	.	.	(3)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	BABY	1	.	.	.	.	.	(2)	.	.	.	.	1	(1)	1	(6)	3	1	.	.	.	.	.	1	.	.	2	.	2	.	(4)	.	.
	LEAF	.	.	.	.	.	.	.	.	.	(4)	1	1	.	1	1	.	1	1	.	.	.	.	.	.	.	.	(2)	.	.	.	.	.
	CEREAL	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	1	.	.	.	.	.	.
	TREE	.	.	.	.	.	.	1	.	.	(2)	1	.	.	.	.	.	2	.	.	.	.	.	1	.	.	(3)	.	.	.	.	.	.
	PURSE	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	3	.	.	.	.	.
	CLOTHES	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	(4)	.	.	.
	PANTS	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	(6)	.	.	.	.
	DOG	.	.	.	.	.	.	(1)	.	.	.	.	.	.	.	(4)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.
	HAMMER	1	.	.	.	.	.	1	1	.	.	2	1	.	.	.	.	.	.	.	2	.	2	2	2	1	.	.	.	.	.	.	1
	CAR	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.

Figure 7. Agreed errors for tests and pretests of chimpanzee Washoe. (IL = within lips; OL = around lips; MF = midface; HD = upper head and side of head; HN = hand; PE = periphery. Conceptual errors are indicated by parentheses.)

	Targets	IL				OL			MF					HD						HN											PE				
	Errors	NUT	CORN	TOOTHBRUSH	SANDWICH	LIPSTICK	BIRD	ICECREAM	BUG	FLOWER	HANKIE	PEEKABOO	SEE	COW	HAT	COMB	CAT	APPLE	PEACH	KEY	CARROT	CHEESE	BALL	BERRY	MEAT	BANANA	OIL	SODAPOP	SHOE	COFFEE	BRUSH	BABY	TREE	DOG	CAR
	NUT	.	.	.	.	.	.	(1)	.	.	.	.	.	.	.	.	.	.	.	.	(1)	.	.	.	.	.	1	.	.	.	.	.	.	.	.
	CORN	.	.	1	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	(1)	.	.	.	.	.	.	.	.	.	.
IL	TOOTHBRUSH	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	(1)	.	.	.	.
	FOOD	.	.	.	(2)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	SANDWICH	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	(1)	.	.	.	(2)	.	.	(6)	.	.	.	.	.	.	.	.	.	.
	SWEET	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	(1)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	RED	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
OL	LIPSTICK	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	(1)	.	.	.	.	.	.	.	.
	BIRD	.	1	.	.	.	.	.	(3)	.	.	.	.	(2)	.	.	(2)	4	.	.	.	.	.	4	.	.	.	.	.	.	.	.	.	(2)	3
	ORANGE	.	.	.	.	.	7	.	.	.	.	.	.	.	.	.	.	.	(1)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	1
	ICECREAM	(1)	.	.		.	.	.	.	.	.	.	1	.	.	.	.	.	(1)	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	1
	BUG	.	.	.	.	.	(1)	1	.	.	.	.	.	(1)	.	.	(4)	.	1	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	1
	FLOWER	.	.	.	.	.	.	.	1	.	.	.	.	.	1	.	1	.	.	1	.	.	.	1	.	1	.	.	.	.	.	1	.	1	1
MF	POTTY	.	.	.	.	.	.	.	2	.	.	.	.	.	.	.	2	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.
	HANKIE	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	PEEKABOO	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1
	SEE	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.
	BLACK	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	COW	1	.	.	.	.	(4)	.	.	.	.	.	.	.	.	.	(6)	.	.	.	.	.	2	.	.	.	.	.	.	.	.	(1)	.	.	.
	HAT	.	.	.	.	.	1	.	1	.	.	.	.	.	.	1	3	.	.	.	.	.	.	.	.	.	.	.	.	.	1	1	.	4	4
	COMB	1	.	.	.	.	.	.	1	.	.	.	.	.	4	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	(4)	3	.	.	2
HD	CAT	.	.	.	.	.	.	.	.	.	.	.	.	(1)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	APPLE	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	PEACH	.	(1)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	(1)	.	.	.	.	.	.	.	.	.
	HEAR	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.
	GIRL	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.		.	.	.	.	.	.	.	.	.
	BLANKET	.	.	.	.	.	.	.	.	.	.	.	.	1	(1)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	KEY	2	.	1	.	1	1	.	.	.	1	1	.	.	.	1	.	2	.	.	3	.	.	.	.	1	1	.	3	.	.	.	.	3	.
	CARROT	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	(1)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	CHEESE	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	(1)	2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	BALL	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	BERRY	.	.	.	.	1	.	.	1	.	.	.	1	1	.	1	.	.	(1)	.	(1)	.	.	.	.	(1)	1	.	.	.	1	.	.	.	.
HN	MEAT	.	(1)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	BANANA	.	.	2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	OIL	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	SODAPOP	(2)	.	.	.	.	.	.	.	.	.	.	.	.	3	.	2	.	.	.	.	(1)	.	.	.	.	2	.	12	.	.	.	.	.	1
	SHOE	.	.	.	.	.	.	.	.	.	.	.	.	.	(1)	.	.	.	.	.	.	.	.	.	.	.	.	7	.	.	.	.	.	.	.
	COFFEE	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.
	BRUSH	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.
	BABY	.	.	.	.	.	.	1	(1)	.	.	.	.	(4)	.	.	(2)	.	2	.	.	.	.	1	.	.	.	.	1	.	1	.	.	(2)	2
PE	TREE	.	.	.	.				.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	DOG	.	.	.	.	.	.	.	.	.	.	.	.	(3)	2	.	(3)	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.
	CAR	.	.	.	.	.	1	.	2	.	1	.	.	.	.	.	2	2	.	.	.	.	1	1	.	1	.	.	1	.	1	1	3	3	.

Figure 8. Agreed errors for tests and pretests of chimpanzee Tatu. (IL = within lips; OL = around lips; MF = midface; HD = upper head and side of head; HN = hand; PE = periphery. Conceptual errors are indicated by parentheses.)

Place Errors

For most of the signs in the vocabularies of Washoe, Moja, Tatu, and Dar, one hand contacted a place on the body or both hands contacted symmetrical places (B. Gardner & Gardner, 1975, Appendix). In the normal conversation of adult signers, contact may

	Targets	IL				OL		MF		HD					HN										PE			
	Errors	NUT	CORN	TOOTHBRUSH	SANDWICH	BIRD	ICECREAM	BUG	FLOWER	HAT	CAT	APPLE	PEACH	COW	CARROT	CHEESE	BALL	BERRY	MEAT	BANANA	OIL	SODAPOP	SHOE	COFFEE	BABY	TREE	DOG	CAR
	DRINK	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	(2)	.	.	.	.
	NUT	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.
IL	CORN	.	.	3	(2)	.	.	1	.	.	.	.	.	.	(2)	.	.	(1)	(1)	.	.	.	.	.	.	.	.	.
	TOOTHBRUSH	.	.	.	.	.	1	.	.	.	.	.	.	.	5	.	.	.	.	3	.	.	.	.	.	.	.	.
	SANDWICH	.	(2)	.	.	.	.	.	.	.	.	.	(1)	.	.	.	.	.	(3)	.	.	.	.	.	.	.	.	.
	BIRD	.	.	.	.	.	2	(5)	6	.	(4)	.	.	(2)	.	.	.	1	.	.	.	.	.	.	.	.	(1)	.
OL	ORANGE	.	.	.	.	2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	ICECREAM	.	.	.	(1)	1	.	.	.	2	.	(2)	.	.	.	.	.	.	.	.	1	(1)	.	.	.	.	.	.
MF	BUG	1	.	1	.	.	.	.	8	.	(1)	.	.	.	.	.	.	4	.	.	.	.	.	.	.	.	.	.
	FLOWER	.	.	.	.	6	.	7	.	1	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.
	HAT	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	COMB	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	CAT	.	1	.	.	(3)	.	(1)	2	.	.	4	6	(13)	.	.	.	.	.	.	.	.	.	.	(2)	.	(7)	1
HD	APPLE	.	.	.	.	1	.	.	.	.	.	.	(2)	1	.	.	.	.	.	.	.	.	.	.	1	.	.	.
	PEACH	.	.	.	.	.	.	.	.	.	1	(2)	.	2	.	.	.	.	.	.	1	.	.	.	1	.	.	.
	COW	.	.	.	.	(1)	.	(1)	.	.	(2)	2	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	LISTEN	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	1	.	.	.
	CARROT	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	(1)	.	.	.	.	.	.	.	.
	CHEESE	.	.	.	.	.	.	.	.	.	.	(1)	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.
	BALL	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.
	BERRY	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
HN	MEAT	.	(2)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	BANANA	.	.	1	.	.	.	.	.	.	.	.	.	.	(2)	.	.	.	.	.	.	.	.	.	.	.	.	.
	OIL	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	3	.	.	.	.	.	.	.	.	.
	SODAPOP	.	.	.	.	.	(6)	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	.	(3)	1	.	.	.
	SHOE	.	.	.	1	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	1	.	.	2	1	.	.	6
	COFFEE	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	BRUSH	.	.	(1)	.	.	.	.	1	.	1	.	.	.	.	.	.	.	.	.	.	1	.	.	.	.	1	.
	BABY	.	.	.	.	(1)	.	.	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	1	.	3
PE	TREE	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	DOG	.	.	.	.	(2)	.	.	.	1	(25)	.	.	(5)	.	.	.	.	.	.	.	.	.	.	(1)	.	.	.
	CAR	.	.	.	.	.	.	.	1	.	.	.	.	.	.	.	.	.	.	.	.	.	1	.	3	.	.	.

Figure 9 Agreed errors for tests and pretests of chimpanzee Dar. (IL = within lips; OL = around lips; MF = midface; HD = upper head and side of head; HN = hand; PE = periphery. Conceptual errors are indicated by parentheses.)

be indicated rather than actually made. The *Dictionary of American Sign Language* (DASL; Stokoe et al., 1976), which is oriented toward adult usage, uses the term *Tab* in a way that largely overlaps with *place* as used here (Stokoe, 1978a, pp. 56–58). The rows and columns of Figures 7, 8, and 9 are arranged according to place of contact, beginning with the mouth, then up the face to the top of the head, down the side of the head to the hands, then along the arms, to the trunk and legs. This sequence of places agrees with Fulton's homunculus (Penfield & Boldrey, 1937) rather than the sequence used in the DASL.

A disproportionately large number of signs in the DASL are assigned to a single Tab designated as the chin and lower face.

Perhaps, more perceptual-motor distinctions can be made per centimeter in this region than in the others. Following Fulton's homunculus, we divided the signs that the DASL places in the lower face region into two places, one within the mouth region (IL for inside the lips) and one outside the mouth region (OL for outside the lips).

In order to assign the signs of all 3 subjects to the same set of places, and to accomplish this in a way that would distribute the errors roughly evenly among the rows and columns of the contingency table, we combined the head and side-of-head regions into a single region (HD, for head region), and we combined all of the signs that would be placed outside the face, head, and hand regions into a single region (PE for peripheral). The six regions that define the cells of Figures 7, 8, and 9 are compatible with the DASL and with Poizner and Lane's (1978) analysis of adult human signing.

Figures 7, 8, and 9 show that a large proportion of the errors fall into the cells that lie along the diagonal of the 6×6 matrices. That is to say, a disproportionate number of errors are correct with respect to place but incorrect with respect to one of the other aspects or dimensions of the target sign. Thus, CAT was a common error for APPLE targets, and vice versa. The sign CAT is a movement that might simulate pulling at imaginary whiskers at the side of the face. The sign APPLE is made by rubbing the knuckles of the fisted hand against the side of the face. BUG was a common error for FLOWER, and vice versa. The sign BUG is made by touching the thumb to the nose. The sign FLOWER is made by bringing the bunched fingertips to the nose.

Those who are unfamiliar with sign languages often suppose that the forms of the signs are highly correlated with their referents. Indeed, in artificial sign languages this is often the case. In the system developed by Paget and Gorman (Crystal & Craig, 1978), for example, all of the signs for animals have similar forms. In natural languages, however, the terms for items in the same conceptual group tend to be distributed widely over the possible gestural or phonetic forms. To show how much relation or lack of relation there was between place and concept in Figures 7–9, we placed in parentheses the cases in which errors belonged to the same conceptual groups as the targets. For this purpose, errors and targets were assigned to groups as in Table 1. Figures 7–9 show that a large proportion of the errors did indeed consist of signs that fall into the same conceptual group as the target signs; for example, DOG for CAT, COMB for BRUSH, and LEAF for FLOWER were prominent errors.

To measure the extent to which correct place determined the errors independently of correct concept, we subtracted the numbers in parentheses from all of the cells and computed a 6×6 chi-square for each of Figures 7, 8, and 9. The chi-squares for errors determined by place independent of concept were 70.1 for Tatu, 125.0 for Dar, and 392.8 for Washoe. With 25 degrees of freedom, a chi-square of 55.0 is required for significance at the .0005 level of significance.

Thus, relations between signs that are based on their gestural form had a highly significant effect on the distribution of errors. According to Hockett (1978, pp. 275–276), evidence of this sort is necessary to establish "duality of patterning." It may be difficult to convey this point to readers who are unfamiliar with ASL, but perhaps this example will help. Warden and Warner (1928) tested the German shepherd dog Fellow, a star of movies and vaudeville. Warden and Warner were concerned with Fellow's ability to understand English to the point where they excluded tests in which Fellow obeyed English instructions combined with visual signals. Fellow's master spoke from behind a screen and instructed him to get one of a set of objects that the experimenters had placed in another room. Correct as well as incorrect trials indicated familiarity with English because Fellow would make errors such as bringing back a *collar* instead of a *dollar*. In the same way, we would argue that only a chimpanzee that had learned ASL would confuse CAT with APPLE or BUG with FLOWER.

Conceptual Errors

The analysis of conceptual errors involves additional problems that were not encountered in analyzing place errors. First, the conceptual relations among vocabulary items are multidimensional. The distribution of CAR errors, for example, suggests that this item fitted better into the Animate group than into the Other (or Miscellaneous Inanimate) group into which it was placed on a priori grounds in constructing Table 1. Second, some conceptual groups yielded very few errors or none at all. To get around these problems, we recast the errors of Figures 7–9 into 3 × 3 tables according to conceptual groups. These groups were Animates, Foods and Drinks, and Other. Errors and targets were assigned to these groups on the basis of their a priori assignment in Table 1.

To measure the extent to which correct concept determined the errors independently of place, we subtracted out all of the cases in which errors were correct as to place; that is, we subtracted all errors that fell into the diagonal cells of Figures 7, 8, and 9 from the 3 × 3 contingency table for concepts. The chi-squares for concept independent of place were 26.9 for Tatu, 50.8 for Dar, and 84.1 for Washoe. With 4 degrees of freedom, a chi-square of 20.0 is required for significance at the .0005 level of confidence.

Moja

Because Moja did not have a series of pretests beyond her initial adaptation sessions, we analyzed only the errors that she made during her single test. Although Moja made more errors during this test than the other subjects made during their tests, she had too few agreed errors to support the same precise statistical analyses. Nevertheless, when her total of 36 agreed test errors are classified in the same way as those of the other subjects, 11 were correct with respect to place only, 14 were correct with respect to concept only, 11 were correct with respect to both, and only 4 of the 36 were correct with respect to neither. Thus, the pattern of Moja's errors was much the same as the pattern for Washoe, Tatu, and Dar.

Indecisive Replies

As noted earlier, some replies contained two or more names for objects. Pairs of signs within such indecisive replies formed roughly the same patterns as those in Figures 7, 8, and 9. Conceptual pairs, such as CAT and DOG or SODAPOP and ICECREAM, and Form pairs, such as CAT and APPLE or BUG and FLOWER, were the most common. Sometimes, pairs were repeated as in CAT DOG CAT DOG or BUG FLOWER BUG FLOWER. Sometimes, replies contained a string of related signs as when Washoe signed CAT BIRD DOG MAN for a picture of a kitten or FLOWER TREE LEAF FLOWER for a picture of daisies. There are several signs that are made by grasping points along the edge of one hand with the thumb and index finger of the active hand; the end of the thumb is grasped in BERRY, the upper edge of the palm in MEAT, the lower edge of the palm in OIL (see B. Gardner & Gardner, 1975, Appendix). In a typical case, Washoe signed OIL BERRY MEAT for a picture of frankfurters, as if the correct sign was on the tip of her fingers. Not only the correct replies but the errors and the very dithering between alternatives were governed by conceptual relations among the referents and phonological relations among the signs of ASL.[4]

Discussion

Observers who were fluent in ASL and familiar with the chimpanzee subjects agreed in their reports of signing. In particular, independent observers agreed with each other under conditions in which their only source of information was the chimpanzee subject. The observers agreed with each other whether the naming was correct or incorrect.

Communication and Information

When Washoe was 27 months old, she made a hole in the flimsy inner wall of her house trailer, located high up in the wall above her bed. Before we repaired the hole, she managed to lose a toy in the hollow space between the inner and outer walls. When the first author arrived that evening,

[4] In using the term *phonology* here, we follow Stokoe (1978a, pp. 43–44) who used *phonology*, *phonetics*, and *phonemics* to refer to analogous aspects of the analysis of spoken and signed languages.

Washoe attracted his attention to an area of the wall down below the hole at the level of her bed, signing OPEN OPEN many times over that area. It was not hard for Allen Gardner to understand what the trouble was and eventually to fish out the toy. When the toy was found, it was exciting to realize that a chimpanzee had used a human language to communicate truly new information. It was not long before situations such as the following became commonplace. Washoe's playground was a suburban garden behind a single-storey, ranch-style home. High in her favorite tree, Washoe was often the first to see cars arriving at the front of the house, and her companions on the ground learned to rely on her to tell them who was coming.

In these vocabulary tests, we verified the field observations under formal experimental conditions. The chimpanzees used ASL to tell the observers things that the observers did not know. This is what was absent in the performances of Clever Hans, the German horse that seemed to do arithmetic by tapping out numbers with his hoof. Not the circus trainers or the cavalry officers, not the veterinarians or the zoo directors, not even the philosophers and the linguists who studied the case could explain how Clever Hans did it. Nevertheless an experimental psychologist, Oskar Pfungst (1911), unraveled the problem with the following test. Pfungst whispered one number into Clever Hans left ear, and Herr von Ost, the trainer, whispered a second number into the horse's right ear. When Clever Hans was the only one who knew the answer, he could not tap out the correct sums. He could not tell his human companions anything that they did not already know.

Since then, controls for "Clever Hans errors" have been standard procedure in comparative psychology.[5] To date, most, if not all, research on human children has been carried out without any such controls. It is as if students of child development believed that whereas horses and chimpanzees may be sensitive to subtle nonverbal communication, it is safe to assume that human children are totally unaffected. A typical example is the frequently cited study of Fraser, Bellugi, and Brown (1963) in which 3-year-old children were shown a number of paired pictures. Each pair illustrated a contrast such as "the dog is biting the cat" versus "the cat is biting the dog." The pictures were untitled, of course, and the children were shown each pair in turn and asked to point to the one that corresponded with the sentence that the observer had just uttered. Fraser et al. commented, "S sometimes pointed quickly, then reflected and corrected himself; the last definite pointing is the one we always scored" (p. 129). As they were in full view of the children, the observers could indicate the correct response by gaze direction alone, or by approving and disapproving facial expressions. With only two possible alternatives, they could indicate the correct reply without uttering a word, without even being aware that they were doing so.

Letting the children change their replies is not the error of Fraser et al.; what is wrong is that the observers knew the correct answer *before* the children replied. By contrast, the observers in the tests reported here never knew what picture was on the screen until *after* the chimpanzees replied. As in the case of Washoe telling us where she lost her toy and who was coming, the chimpanzees in these tests knew something that the observers did not know. They had information to communicate. The purpose of the Clever Hans control is not to satisfy some abstract principle of comparative psychology; the purpose is to prove that information has been communicated.

Communication and Concepts

In order to make sure that the signs referred to conceptual categories, all of the trials in Tests 1 and 2 were first trials; that is, each slide was shown to the subject for the first time on the one and only trial in which it was presented to that subject. The typical range of exemplars in each category is illustrated in Figures 3, 4, 5, and 6. Specific stimulus values varied, as they do in natural language categories. Most human beings would agree that the exemplars in each set belong together; apparently, Washoe, Moja, Tatu, and Dar agreed also.

---

[5] In modern times, a striking exception to this rule has been the work of Terrace (1979) with the chimpanzee Nim which included no controls for Clever Hans errors, whatsoever.

When teaching a new sign, we usually began with a particular object—a particular toy for BALL, a particular shoe for SHOE. At first, especially with very young subjects, there would be very few balls and very few shoes. The same situation is common in human nursery life. Early in Project Washoe we worried that the signs might become too closely associated with their initial referents. It turned out that this was no more a problem for Washoe or our other subjects than it is for children. The chimpanzees transferred the signs they had learned for a few balls, shoes, flowers, or cats to the full range of the categories whenever they found them and however represented, as if they divided the world into conceptual categories just as we do.

We were surprised by the apparently spontaneous transfer of MEAT from cooked to fresh meats. Like human children, they ate meat, but unlike human children, nearly all the meat they saw was intended for their individual consumption, so it came in small portions and was mostly pre-cooked. The chimpanzees seemed to be able to identify the pictures of large, cooked roasts that often appear in women's magazines, but we never took them to markets to see large portions of fresh meat.

One day, when 5½-year-old Moja was in the Gardner kitchen for her weekly brunch, she looked at a large cut of fresh beef on a spit waiting to be roasted for a later Gardner meal, and signed MEAT. This was a surprise to us. No one recalled any earlier incident in which fresh meat had been named for her, or for any of the other subjects. It was unlikely that anyone would have done so, because we all assumed that fresh cuts of meat would be meaningless and uninteresting to these chimpanzees. After that, we asked Moja as well as Tatu and Dar about many kinds of fresh meat, and they named these as readily as they named familiar portions of cooked meat. As Figure 5 shows, Tatu's and Dar's exemplars for MEAT included uncooked steaks and chops fresh from the supermarket.

We probably did not teach the chimpanzees to include both fresh and cooked meat in the MEAT category. The transfer seems to have been spontaneous. But how does anyone teach a concept? We also taught them signs for colors, but this would have been impossible with color-blind chimpanzees. By definition, unless the learners can see the difference between red and green in the first place, no teacher, however skilled or however talented, could possibly teach them to use words or signs to name these colors correctly. By the same token, no one could teach chimpanzees or children to use signs or words to distinguish between photographs of meats and photographs of nonmeats unless (a) there is some visual difference between meats and nonmeats and (b) their visual apparatus is sensitive to that difference. The same argument applies with equal force to all of the conceptual categories in the present investigation.

Apart from the theoretical questions that it raises, the apparently spontaneous transfer of MEAT illustrates certain critical features of our procedure. First, the initial observation depended on the enriched, open-ended conditions of cross-fostering. Second, the continual presence of open-minded, professional, fluent observers, who were also well-acquainted with the chimpanzees, increased the chances that the unexpected would, nevertheless, be reported. Third, we confirmed these observations with the same subject and with different subjects, under similar informal conditions. Finally, we confirmed the informal observations under systematic test conditions, conditions precise enough to show that errors as well as correct replies were governed by conceptual categories.

No animal, human or nonhuman, steps twice into the same perceptual stream. A young monkey that picks a ripe mango in a tree must learn something general about ripe mangoes, because it will certainly never see that particular mango again. A young lion that stalks and kills an impala must learn something general about hunting impalas, because it will certainly never meet that particular impala again. Whatever may be learned about a particular object or event, the most valuable things that are learned have to do with variable aspects of general stimulus classes. What the psychologist must explain is how past experience transfers to new situations, because this is what happens most of the time.

So much experimentation has been limited to precisely repeated stimulus objects or to objects that vary only in simple dimensions, such as color and size, that it would be easy to form the impression that the conceptual abilities of nonhuman beings are severely limited. But, there are notable exceptions. Hayes and Hayes (1953) with chimpanzee Viki, Hicks (1956) and Sands, Lincoln and Wright (1982) with monkeys, and Herrnstein, Loveland, and Cable (1976) with pigeons, for example, have all demonstrated that nonhuman beings can use natural language categories when they are tested with suitably varied stimulus material.

Significant variation among exemplars and testing with true first trials are essential to the definition of natural language categories. More concerned with theoretical definitions of language than with conceptual behavior, the Rumbaughs (Essock, Gill, & Rumbaugh, 1977; Gill & Rumbaugh, 1974; Savage-Rumbaugh, Pate, Lawson, Smith, & Rosenbaum, 1983) have administered hundreds of trials of training and testing with identical exemplars or with minimally varied exemplars. To be sure, the Rumbaughs have concentrated on the arbitrariness of what they call "lexigrams" and their use in arbitrarily fixed sequences. It seems likely that the Rumbaugh subjects could use natural language categories, given the opportunity to do so.

In 4 years of work with the chimpanzee Nim, Terrace and his associates (Terrace, 1979; Terrace, Petitto, Sanders, & Bever, 1980) never attempted any experimental analysis of reference. Their work is unique in this field in that they never administered any systematic tests at all. Since the time of Kellogg and Kellogg (1933), it is the first study in this field that was entirely restricted to adventitous naturalistic observation.

Communication and Language

Although concern with grammar has occupied so much of the efforts of developmental psycholinguists, in our view it would be a mistake for psychobiologists to neglect method and theory in the study of reference. For if the development of human verbal behavior requires any significant expenditure of biological resources, then it must confer selective advantages on its possessors. In order to confer any selective advantage, however, a biological trait must operate on the world in some way; it must be instrumental in obtaining benefit or avoiding harm. If clarifying one's ideas confers selective advantage, it must be because in some way clarified ideas provide superior means for operating in the biological world. As for establishing social relations, a system of displays and cries is sufficient to maintain group cohesiveness in most animals. The selective advantage of a wider variety of signals would seem to be the communication of more information. But, unless verbal behavior refers to objects and events in the external world, it cannot communicate information and it cannot have any selective advantage. From this point of view, reference is the Darwinian function of verbal behavior, and the function of grammar or structure in verbal behavior must be to enlarge the scope and increase the precision of reference.

In the naturally occurring languages of the world, the pairing of words and signs with conceptual categories is arbitrary. This is amply demonstrated by the mutual unintelligibility of languages and the well-documented history of shifts in forms and usage (Laird, 1981; Lehmann, 1973). International meetings, such as the World Federation of the Deaf, require simultaneous translation, often as many as five simultaneous translators (Battison & Jordan, 1976; Jordan & Battison, 1976), and historical shifts in ASL have been documented by Frishberg (1975) and Woodward (1976). Washoe, Moja, Tatu, and Dar, if they had been human children—if they had been angels for that matter—could succeed in these vocabulary tests only by associating the signs of ASL with their referents. Angels may have other ways of associating responses with stimuli, but children and chimpanzees must learn arbitrary associations. Thus, to the extent that the communication of information depends on the arbitrary pairing of terms with conceptual categories, then that biological function of a natural language depends on the rote learning of paired associates.

Communication and Motive

Because no extrinsic reward seems to be necessary, many have claimed that human children learn to speak as if they had an inborn, species-specific motive to communicate (e.g., Sugarman, 1983, p. 496). Ethologists and comparative psychologists have shown, however, that any general account of motivation must allow for a variety of motives rather than a few basic ones, such as hunger and thirst, which give rise to the rest through a process of conditioning. Moreover, other sources of motivation can be more powerful determinants of behavior than hunger and thirst. Harlow's (1958) experiments on contact comfort come immediately to mind. It is also clear now that many nonhuman species behave as if they had a need to communicate, and such a need is by no means uniquely human (Tinbergen, 1953). Needs such as contact comfort and communication have obvious selective advantages. To the modern mind, the existence of many unlearned motives is more compatible with Darwinism than the laborious process of conditioning with food and drink that was formerly posited.

Chimpanzees are among the many species that behave as if they have a strong need to communicate (Van Lawick-Goodall, 1968). Captive chimpanzees are similar to wild chimpanzees in this respect (Kellogg, 1968) unless their conditions of captivity are so severe that normal behavior is suppressed. In the Reno laboratory, when we did introduce extrinsic rewards for signing as, for example, when we rewarded Tatu and Dar with treats for obedient test-taking behavior, the extrinsic rewards usually interfered and had to be discontinued.

The following example taken from Hayes and Nissen (1971) is typical of cross-fostered chimpanzees.

> One hot summer day [Viki] brought a magazine illustration of a glass of iced tea to a human friend. Tapping it, she said "Cup! Cup!" and ran to the refrigerator, pulling him along with her. It occurred to us that pictures might be used to signify needs more explicitly than words . . . .
>
> A set of cards was prepared showing magazine illustrations in natural color of those things she solicited most frequently. [For four days Viki consistently used the picture-cards for requests, but on the fifth day] . . . suddenly she acted as if imposed upon. She had to be coaxed to cooperate and then used the pictures in a completely random way.
>
> [After seven months of erratic performance] . . . the technqiue which had seemed so promising was dropped, pending revision. Spring weather, plus a new car, gave Viki a wanderlust so that no matter what situation sent her to the picture-communication pack, when she came upon a car picture she made happy noises and prepared to go for a ride. We eliminated all car pictures from the pack, but it was too late. Long afterwards Viki was tearing pictures of automobiles from magazines and offering them as tickets for rides. (pp. 107–108)

Most, if not all, of the techniques used by the Rumbaughs and by Terrace can be described as teaching the chimpanzee elaborate ways of begging for food and other commodities (Essock et al., 1977; Gill & Rumbaugh, 1974; Savage-Rumbaugh et al., 1983; Terrace, 1979; Terrace et al., 1980). At the same time, working with human children, Lepper, Greene, and Nisbett (1973) and Levine and Fasnacht (1974) demonstrated that the heavy-handed application of extrinsic rewards impairs the performance of intrinsically motivated tasks, such as drawing. Heavy reliance on extrinsic rewards probably has a similarly negative effect on the performance of chimpanzees. Characteristically, those who have relied most heavily on extrinsic rewards have been those who most insistently claimed that chimpanzees lack intrinsic motivation to communicate (Savage-Rumbaugh & Rumbaugh, 1978; Savage-Rumbaugh et al., 1983, pp. 462, 485–486; Terrace, 1979, pp. 221–224; Terrace et al., 1980, pp. 438–440).

Obviously, as their verbal skills improved, Washoe, Moja, Tatu, and Dar were more successful in making their wants known and, presumably, more successful in getting what they wanted. But certainly the same can be said for human children—or human adults, for that matter. Also, those serving as O_1 often showed their approval or disapproval by smiling or frowning, by nodding or shaking their heads, and by praising the chimpanzees in ASL. This was carried over from the cross-fostering regime that was maintained at all times, and it is the way human adults normally respond to the verbal behavior of human children.

As an objective measuring instrument or an empirical tool for studying the early vocabulary, the procedures reported here are clearly superior to the highly interpretive, naturalistic methods that have so far dominated studies of the early vocabulary of young children (Nelson, Rescorla, Gruendel, & Benedict, 1978). An important contribution of comparative psychology has always been the application of rigorous principles of experimental method to fundamental problems of developmental psychology.

Considering the immaturity of Washoe, Moja, Tatu, and Dar at the time of these vocabulary tests, we cannot use these results to estimate the limits of their ability. Chimpanzees begin to lose their milk teeth when they are 5 or 6 years old. Under natural conditions in Africa, infants are not weaned until they are 4 or 5 years old; they usually live with their mothers until they are 7 and often continue to live with their mothers until they are 10 or 11. The youngest wild chimpanzee mother at the Gombe Stream was 12 years old when her first baby was born (Van Lawick-Goodall, 1973). Although the life span of wild chimpanzees is still unknown, we do know that captive chimpanzees can remain vigorously and intelligently alive for more than 50 years (Maple & Cone, 1981). It is clear that a 5-year-old chimpanzee is far from realizing the limits of chimpanzee intelligence and the full benefits of cross-fostering. The strongest conclusion that we can make about the results so far is that there is much more to be discovered about the continuity between human and nonhuman intelligence.

References

Battison, R., & Jordan, I. K. (1976). Cross-cultural communication with foreign signers: Fact and fancy. *Sign Language Studies, 10*, 53-68.

Brown, J. L. (1975). *The evolution of behavior.* New York: Norton

Brown, R. (1970). The first sentences of child and chimpanzee. In R. Brown (Ed.), *Selected psycholinguistic papers* (pp. 208-281). New York: Macmillan.

Crystal, D., & Craig, E (1978). Contrived sign language. In I. M. Schlesinger & L. Namir (Eds.), *Sign language of the deaf* (pp. 141-168). New York: Academic Press.

Diaconis, P. (1978). Statistical problems in ESP research. *Science, 201*, 131-136.

Essock, S. M., Gill, T. V , & Rumbaugh, D. M. (1977). Language relevant object- and color-naming tasks. In D. M. Rumbaugh (Ed.), *Language learning by a chimpanzee* (pp. 193-206). New York: Academic Press.

Fant, L. (1977). *Sign language.* Northridge, CA: Joyce Media, Inc.

Fouts, R S. (1972). Use of guidance in teaching sign language to a chimpanzee. *Journal of Comparative and Physiological Psychology, 80*, 515-522.

Fraser, C., Bellugi, U., & Brown, R. (1963). Control of grammar in imitation, comprehension and production. *Journal of Verbal Learning and Verbal Behavior, 2*, 121-135.

Frishberg, N. (1975). Arbitrariness and iconicity: Historical change in American Sign Language. *Language, 51*, 696-719.

Gardner, B. T., & Gardner, R. A (1971). Two-way communication with an infant chimpanzee. In A. Schrier & F. Stollnitz (Eds.), *Behavior of nonhuman primates* (Vol. 4, pp. 117-184). New York: Academic Press.

Gardner, B. T., & Gardner, R. A. (1974) Comparing the early utterances of child and chimpanzee. In A. Pick (Ed.), *Minnesota symposium on child psychology* (Vol. 8, pp. 3-23). Minneapolis: University of Minnesota Press.

Gardner, B. T., & Gardner, R A. (1975). Evidence for sentence constituents in the early utterances of child and chimpanzee. *Journal of Experimental Psychology: General, 104*, 244-267.

Gardner, B. T., & Gardner, R. A. (1980). Two comparative psychologists look at language acquisition. In K. E. Nelson (Ed.), *Children's language* (Vol. 2, pp. 331-369). New York: Gardner Press.

Gardner, R. A., & Gardner, B. T. (1969). Teaching sign language to a chimpanzee. *Science, 165*, 664-672.

Gardner, R. A., & Gardner, B. T. (Producers and Directors) (1973). *Teaching sign language to the chimpanzee Washoe* [Film]. State College, PA: Psychological Cinema Register.

Gardner, R. A., & Gardner, B. T. (1974). Teaching sign language to the chimpanzee, Washoe. *Bulletin D'Audio Phonologie, 4*, 145-173.

Gardner, R. A., & Gardner, B. T. (1978). Comparative psychology and language acquisition. *Annals of the New York Academy of Sciences, 309*, 37-76.

Gill, T. V., & Rumbaugh, D. M. (1974). Mastery of naming skills by a chimpanzee. *Journal of Human Evolution, 3*, 483-492.

Harlow, H. F. (1958). The nature of love. *American Psychologist, 13*, 673-685.

Hayes, K. J., & Hayes, C. (1951). The intellectual development of a home-raised chimpanzee. *Proceedings of the American Philosophical Society, 95*, 105-109.

Hayes,K. J., & Hayes, C. (1953). Picture perception in a home-raised chimpanzee. *Journal of Comparative and Physiological Psychology, 46*, 470-474.

Hayes, K. J., & Nissen, C. H. (1971). Higher mental functions of a home-raised chimpanzee. In A. M. Schrier & F. Stollnitz (Eds.), *Behavior of nonhuman primates* (Vol. 4, pp. 59-115). New York: Academic Press.

Herrnstein, R. J., Loveland, D. H., & Cable, C. (1976). Natural concepts in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 2*, 285-302.

Hewes, G. W. (1973). Pongid capacity for language acquisition: An evaluation of recent studies. *Symposia of the Fourth International Congress of Pri-*

matology, 1, 124–143.
Hicks, L H. (1956). An analysis of number-concept formation in the rhesus monkey. *Journal of Comparative and Physiological Psychology, 49*, 212–218.
Hockett, C. F. (1978). In search of Jove's brow. *American Speech, 53*, 243–314.
Jordan, I K., & Battison, R. (1976). A referential communication experiment with foreign sign languages. *Sign Language Studies, 10*, 69–80.
Kellogg, W. N. (1968). Communication and language in the home-raised chimpanzee. *Science, 162*, 423–427
Kellogg, W. N , & Kellogg, L. A.(1933). *The ape and the child*. New York: Hafner.
Laird, C. (1981). *The word*. New York: Simon & Schuster.
Lehmann, W. P. (1973). *Historical linguistics: An introduction* (2nd ed.). New York: Holt, Rinehart & Winston.
Lepper, M R., Greene, D., & Nisbett, R. E (1973). Undermining children's intrinsic interest with extrinsic reward· A case of the "overjustification" hypothesis. *Journal of Personality and Social Psychology, 28*, 129–137.
Levine, F. & Fasnacht, G. (1974) Token rewards may lead to token learning *American Psychologist, 29*, 816–820.
Lieberman, P. (1984) *The biology and evolution of language*. Cambridge, MA: Harvard University Press.
Maple, T. L., & Cone, S. G. (1981). Aged apes at the Yerkes Regional Primate Research Center. *Laboratory Primate Newsletter, 20*, 10–12
Nelson, K., Rescorla, L., Gruendel, J., & Benedict, H (1978). Early lexicons: What do they mean?. *Child Development, 49*, 960–968.
Patterson, J. C. (1983, February). *Computational formulas for use in card-guessing with information*. Paper presented at the Mid-South Colloquium in Mathematical Sciences, Memphis, TN.
Penfield, W. G , & Boldrey, E (1937). Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation *Brain, 60*, 389–443
Pfungst, O. (1911). *Clever Hans* (C. L. Rahn, trans.) New York: Holt
Poizner, H., & Lane, H (1978). Discrimination of location in American Sign Language In P. Siple (Ed.), *Understanding language through sign language research* (pp 271–287). New York: Academic Press.
Read, R. C. (1962). Card-guessing with information—A problem in probability. *American Mathematical Monthly, 69*, 506–511.
Rosch, E., & Lloyd, B. B. (Eds.) (1978). *Cognition and categorization*. Hillsdale, NJ: Erlbaum.
Saltz, E., Dixon, D., Klein, S., & Becker, G. (1977). Studies of natural language concepts: III. Concept overdiscrimination in comprehension between two and four years of age. *Child Development, 48*, 1682–1685.
Saltz, E., Soller, E., & Sigel, I. (1972). The development of natural language concepts. *Child Development, 43*, 1191–1202.
Sands, S. F., Lincoln, C. E., & Wright, A. A. (1982). Pictorial similarity judgments and the organization of visual memory in the rhesus monkey. *Journal of Experimental Psychology: General, 111*, 369–389.
Savage-Rumbaugh, E. S., Pate, J. L., Lawson, J., Smith, S. T., & Rosenbaum S. (1983). Can a chimpanzee make a statement? *Journal of Experimental Psychology: General, 112*, 457–492.
Savage-Rumbaugh, E. S., & Rumbaugh, D. M. (1978). Symbolization, language, and chimpanzees: A theoretical reevaluation based on initial language acquisition processes in four young *Pan troglodytes Brain and Language, 6*, 265–300.
Stokoe, W. C. (1978a). *Sign language structure* (rev. ed.). Silver Spring, MD: Linstok Press.
Stokoe, W. C. (1978b). Sign language versus spoken language. *Sign Language Studies, 18*, 69–90.
Stokoe, W. C. (1983) Apes who sign and critics who don't. In H. T. Wilder & J de Luce (Eds.), *Language in primates*. Bloomington: Indiana University Press.
Stokoe, W. C., Casterline, D. C., & Croneberg, C. G. (1976). *A dictionary of American Sign Language on linguistic principles*. Silver Spring, MD: Linstok Press.
Sugarman, S. (1983). Why talk? *Journal of Experimental Psychology: General, 112*, 493–497.
Terrace, H. S. (1979). *Nim*. New York: Knopf.
Terrace, H. S., Pettito, L., Sanders, R. J., & Bever, T. G. (1980). On the grammatical capacity of apes. In K. E. Nelson (Ed.), *Children's language* (*Vol.* 2, pp. 371–495). New York. Gardner Press.
Tinbergen, N. (1953). *Social behaviour in animals*. New York: Wiley.
Umiker-Sebeok, J., & Sebeok, T. A. (1980). Questioning apes. In T. A. Sebeok & J. Umiker-Sebeok (Eds). *Speaking of apes* (pp. 1–59). New York: Plenum Press
Van Cantfort, T. E., & Rimpau, J. B. (1982). Sign language studies with children and chimpanzees. *Sign Languages Studies, 34*, 15–72.
Van Lawick-Goodall, J. (1968). A preliminary report on expressive movements and communication in the Gombe Stream chimpanzees. In P. C. Jay (Ed.), *Primates: Studies in adaptation and variability* (pp. 313–374). New York: Holt, Rinehart & Winston.
Van Lawick-Goodall, J. (1973). The behavior of chimpanzees in their natural habitat. *American Journal of Psychiatry, 130*, 1–11.
Warden, C. J., & Warner, L. H. (1928). The sensory capacities and intelligence of dogs with a report on the ability of the noted dog "Fellow" to respond to verbal stimuli. *Quarterly Review of Biology, 3*, 1–28.
Watson, D. O. (1973). *Talk with your hands*. Menasha, WI: George Banta.
Woodward, J. C. (1976). Signs of change: Historical variation in American Sign Language. *Sign Language Studies, 10*, 81–94.

Received November 4, 1983
Revision received July 30, 1984 ■