Journal of Scientific Exploration, Vol. 17, No. 4, pp. 601–616, 2003 0892-3310/03

Testing a Language-Using Parrot for Telepathy

RUPERT SHELDRAKE

20 Willow Road, London NW3 1TJ, UK
e-mail: ars@dircon.co.uk

AIMÉE MORGANA

PO Box 625, New York, NY 10031

Abstract—Aimée Morgana noticed that her language-using African Grey parrot, N'kisi, often seemed to respond to her thoughts and intentions in a seemingly telepathic manner. We set up a series of trials to test whether this apparent telepathic ability would be expressed in formal tests in which Aimée and the parrot were in different rooms, on different floors, under conditions in which the parrot could receive no sensory information from Aimée or from anyone else. During these trials, Aimée and the parrot were both videotaped continuously. At the beginning of each trial, Aimée opened a numbered sealed envelope containing a photograph, and then looked at it for two minutes. These photographs corresponded to a prespecified list of key words in N'kisi's vocabulary, and were selected and randomized in advance by a third party. We conducted a total of 147 two-minute trials. The recordings of N'kisi during these trials were transcribed blind by three independent transcribers. Their transcripts were generally in good agreement. Using a majority scoring method, in which at least two of the three transcribers were in agreement, N'kisi said one or more of the key words in 71 trials. He scored 23 hits: the key words he said corresponded to the target pictures. In a Randomized Permutation Analysis (RPA), there were as many or more hits than N'kisi actually scored in only 5 out of 20,000 random permutations, giving a p value of 5/20,000 or 0.00025. In a Bootstrap Resampling Analysis (BRA), only 4 out of 20,000 permutations equaled or exceeded N'kisi's actual score ($p = 0.0002$). Both by the RPA and BRA, the mean number of hits expected by chance was 12, with a standard deviation of 3. N'kisi repeated key words more when they were hits than when they were misses. These findings are consistent with the hypothesis that N'kisi was reacting telepathically to Aimée's mental activity.

Keywords: telepathy—interspecies communication—parrot—language-using animals

Introduction

Until the 1980s, within academic science it was generally assumed that parrots were mere mimics, "parroting" words with no understanding. Most scientific studies of human-to-animal linguistic communication were carried out with primates, using sign language (e.g., Patterson & Linden, 1981; Fouts, 1997).

In 1977, Irene Pepperberg began training and testing an African Grey parrot, Alex, and subsequently succeeded in showing that Alex and other parrots can use language meaningfully. Over 20 years of training, Alex acquired a vocabulary of more than 200 words, and Pepperberg established that Alex was capable of abstraction and of using language referentially. For example, he can grasp such concepts as ''present'' and ''absent'' and use words for colors appropriately, whatever the shape of the colored object (Pepperberg, 1999). Pepperberg and her colleagues have shown that parrots, although literally bird-brained, rival primates in their ability to use language meaningfully.

Inspired by seeing Alex on television, in 1997 Aimée Morgana began training a young male African Grey parrot, N'kisi (pronounced ''in-key-see''), in the use of language. She did so by teaching him as if he were a human child, starting when he was 5 months old. She used two teaching techniques known as ''sentence frames'' and ''cognitive mapping.'' In sentence frames, words were taught by repeating them in various sentences such as, ''Want some water? Look, I have some water.'' Cognitive mapping reinforced meanings that might not yet be fully understood. For example, if N'kisi said ''water,'' Aimée would show him a glass of water. By the time he was 5 years old, he had a contextual vocabulary of more than 700 words. He apparently understood the meanings of words and used his language skills to make relevant comments. He ordinarily spoke in grammatical sentences, and by January 2002, Aimée had recorded more than 7,000 original sentences.

Although Aimée's primary interest was in N'kisi's use of language, she soon noticed that he said things that seemed to refer to her own thoughts and intentions. He did the same with her husband, Hana. After reading about Rupert Sheldrake's research on telepathy in animals (Sheldrake, 1999), in January 2000 Aimée contacted Sheldrake and summarized some of her observations. At the same time, she began keeping a detailed log of seemingly telepathic incidents, and has continued to do so. By January 2002, she had recorded 630 such incidents. Here are a few examples:

''I was thinking of calling Rob, and picked up the phone to do so, and N'kisi said, 'Hi, Rob,' as I had the phone in my hand and was moving toward the Rolodex to look up his number.''

''We were watching the end credits of a Jackie Chan movie, edited to a musical soundtrack. There was an image of [Chan] lying on his back on a girder way up on a tall skyscraper. It was scary due to the height, and N'kisi said, 'Don't fall down.' Then the movie cut to a commercial with a musical soundtrack, and as an image of a car appeared, N'kisi said, 'There's my car.' (N'kisi's cage was at the other end of the room, and behind the TV. He could not see the screen and there were no sources of reflection.)

''I read the phrase 'The blacker the berry, the sweeter the juice'; N'kisi said 'That's called black' at the same instant.''

''I was in a room on a different floor, but I could hear him. I was looking at a deck of cards with individual pictures, and stopped at an image of a purple car. I was thinking it was an amazing shade of purple. Upstairs he said at that instant, 'Oh wow, look at the pretty purple.'''

Of all the various incidents, perhaps the most remarkable occurred when N'kisi interrupted Aimée's dreams. (He usually slept by her bed.) For example:

"I was dreaming that I was working with the audio tape deck. N'kisi, sleeping by my head, said out loud, 'You gotta push the button,' as I was doing exactly that in my dream. His speech woke me up."

On another occasion,

"I was on the couch napping, and I dreamed I was in the bathroom holding a brown dropper medicine bottle. N'kisi woke me up by saying, 'See, that's a bottle.'"

In April 2000, we met for the first time at Aimée's home in Manhattan, New York. Together we set up a simple test that replicated a situation in which N'kisi had appeared to demonstrate telepathy spontaneously. We went to another room, where N'kisi could not see what we were doing, and Aimée looked through a pack of cards with various pictures on them. After looking at several other cards, she held up a picture that showed a girl and looked at it intently. As she did so, we heard N'kisi say with unmistakable clarity, "That's a girl." Since we were in a different room, and had not spoken about the image, it seemed very unlikely that clues had been transmitted through any of the normal sensory channels.

Clearly, it was important to try to test this apparent telepathic communication in controlled experiments. We developed a procedure that could work fairly naturally in N'kisi's familiar environment. Aimée had noticed that N'kisi seemed to respond to moments of discovery, as if he "surfed the leading edge" of her consciousness. Therefore, methods of testing for telepathy that used repetitive images, such as playing cards or Zener cards, were not likely to work. In order to preserve an element of surprise, we designed an experiment in which Aimée was filmed as she opened sealed envelopes one at a time, each containing a different photograph. Meanwhile N'kisi was alone in a different room, unable to see or hear Aimée, and he was filmed continuously to record his behavior and speech.

Methods

Selection of Images

The first step in this experimental procedure was the preparation of a list of key words that were part of N'kisi's vocabulary and could be represented by visual images. From his vocabulary list, there were 30 such words, such as "phone," "flower," and "bottle." Another person, Evan Izer, unconnected with the experiment in any other way, selected 167 different photographs on the basis of this list from a stock image supplier, the PhotoDisc Resource Collection. He sealed prints of these images in thick brown opaque envelopes, one image per envelope, and randomized their order by shuffling them thoroughly before numbering the envelopes. Of these 167 sealed envelopes, we used 20 in preliminary tests to work out a standard procedure, leaving 147 for the main experiment.

No one had any way of knowing in advance what photograph Aimée would be looking at in any given trial. Out of the initial list of 30 key words, photographs corresponding to 20 of these words were included in the test. Appropriate

images for the remaining key words could not be found in the photo source. The experiment was based on the 20 key words for which images were available, but one of these words, "camera," had to be eliminated because N'kisi used it so frequently to comment on the cameras used in the tests themselves, as explained below. Thus, the analysis of results was based on 19 key words.

Test procedure

During the tests, N'kisi remained in his cage in Aimée's apartment in Manhattan, New York. There was no one in the room with him. Meanwhile, Aimée went to a separate enclosed room on a different floor. N'kisi could not see or hear her, and in any case, Aimée said nothing, as confirmed by the audio track recorded on the camera that filmed her continuously. The distance between Aimée and N'kisi was about 55 feet. Aimée could hear N'kisi through a wireless baby monitor, which she used to gain "feedback" to help her to adjust her mental state as image sender.

Both Aimée and N'kisi were filmed continuously throughout the test sessions by two synchronized cameras on time-coded videotape. The cameras were mounted on tripods and ran continuously without interruption throughout each session. N'kisi was also recorded continuously on a separate audio tape recorder.

There were 30 test sessions, in most of which there were 5 two-minute trials. In 3 sessions, when Aimée had less time available, there were 4 two-minute trials. The total number of trials was determined in advance by the number of sealed image-containing envelopes available, namely 147. At the beginning of each trial, Aimée set a timer so that it would beep after two minutes, then opened the next numbered sealed envelope containing a photograph, which she looked at for the rest of the trial period. The end of the period was signaled by beeps from the timer. N'kisi was free to say whatever he wished during the trial period.

Out of the 147 photographs, 1 was so obscure that Aimée could not make sense of it when she opened the envelope, and so she discarded it straight away, moving on to the next envelope. Ten others were of images that corresponded to none of the pre-selected key words, for example a picture of ice cream, and were eliminated later. Four had to be eliminated because they involved pictures of cameras; and 1 trial was interrupted by a caller. This left a total of 131 trials for analysis.

We have compiled a brief videotaped documentary of some of the trials, showing simultaneous films of Aimée and N'kisi side-by-side in a split-screen format.

Transcription of the Tapes

Three people independently transcribed the audio tapes from each test session, writing what N'kisi said. Two of these people were in the United States: Anna Yamamoto in New York, and Betty Killa in Arizona. The other transcriber, Pam

Smart, was in England. These three people did not know what pictures Aimée had been looking at, nor know any other details of the tests.

Aimée herself also made a set of transcripts, using the sound track on the videotape rather than the audiotape. On her transcripts, she carefully noted when each trial began (i.e., the point at which the envelope was opened and the image was first visible). She also noted when the trial ended (i.e., when the beeper signaled two minutes had passed and the image was put down). The division of the blind transcripts into portions corresponding to the two-minute trial periods was carried out on the basis of these transcripts derived from the video sound track.

Tabulation of the Results

There was good agreement between the three "blind" transcribers. However, in a few cases, one or more of them missed some words recorded by one or both of the others, or heard them differently. This was most often the case with the English transcriber, Pam Smart, who was not familiar with the American accent in which N'kisi spoke.

To find out how much difference these disagreements made, the data were tabulated in three ways: first, by using the data in which all transcribers were in agreement; second, by taking a majority verdict; and third, by including words recorded by only one out of the three.

Aimée's own transcripts were not used in the scoring process because they were not carried out blind. Nevertheless, her transcripts were in good agreement with those of the blind transcribers, and in almost every case in which only one American transcriber recorded a key word, Aimée had also recorded the same word.

Only trials in which N'kisi said one or more of the key words were included in the analysis of the data, because only in such trials could N'kisi have scored a hit or a miss. Trials in which he said nothing or used words that were not on the list of prespecified key words were not included.

In some trials, N'kisi repeated a given key word. For example, in one trial N'kisi said "phone" three times, and in another he said "flower" ten times, and in the tabulation of data the numbers of times he said these words are shown in parentheses as: phone (3); flower (10). For most of the statistical analyses, repetitions were ignored, but in one analysis the numbers of words that were said more than once in a given trial were compared statistically with those said only once for both hits and misses.

For each trial, the key word or words represented in the photograph were tabulated. Some images had only one key word, but others had two or more. For example, a picture of a couple hugging in a pool of water involved two key words, "water" and "hug."

One of the 20 key words, "camera," had to be eliminated because N'kisi frequently said "camera" when he saw Aimée switching on the cameras prior to

a test session and while the cameras were in use during the tests. Consequently the high degree of ''noise'' associated with this word meant that any ''signal'' would be swamped. This left a total of 19 key words for the analysis, as shown in Table 3.

If N'kisi said a key word that did not correspond to the photograph, that was counted as a miss, and if he said a key word corresponding to the photograph, that was a hit. Thus, for example, with an image of someone speaking on the phone, with no flowers in the picture, if he said ''phone'' and ''flower'' he scored one hit and one miss.

Statistical Analysis

The data were analysed independently by Jan van Bolhuis, assistant professor of statistics at the Free University of Amsterdam, Holland. He used three methods. The first two ignored repetitions of key words by N'kisi during a given trial. The third took repetitions into account.

1. In a Randomized Permutation Analysis (RPA), for all trials in which N'kisi said at least one key word, the key words uttered by N'kisi were combined with all the pictures in 20,000 different random permutations. (For a description of this kind of RPA procedure, see Efron and Tibshirani, 1998.) In this analysis, repetitions of a key word during a given trial were ignored. For example, if he said ''flower'' once or several times during a given trial it was counted as a single response. In this way, we simplified the analysis but lost information.

   Using the majority scoring method, with data as shown in Table 2, all 117 key words said by N'kisi were arranged in a random order, from 1 to 117, and then assigned to the images as shown in Table 2. Thus, the first four words were combined with image 1, the fifth word with image 2, and so on, until the word 117 was combined with the last image. Then this process was repeated over and over again with different random permutations of the 117 key words N'kisi said, until 20,000 randomized permutations of key words had been combined with the images. The assignment of randomly permutated key words to images followed the pattern actually observed, for example with four words assigned to image 1, one word to image 2, and so on.

   A complication in this analysis was that in many of the randomized permutations, in cases where there was more than one word assigned to a given image, some key words were assigned twice or more to a given image. Because this analysis did not include repetitions of a key word in a given trial, all permutations that assigned a given key word to a particular picture more than once were excluded from the analysis. Thus, the 20,000 combinations included in this analysis did not include any repetitions of a key word in a given trial.

   This analysis was carried out using a computer program specially written by Jan van Bolhuis for this purpose. These computations were

extremely time-consuming, and took many days of computer time on a PC (personal computer).

For each of the 20,000 random permutations, the computer counted the number of hits, meaning the number of times that the randomly assigned key words corresponded to images. This enabled the mean number of hits expected by chance to be calculated, together with the standard deviation.

The probability (p-value) of N'kisi's actual score arising by chance was estimated on the basis of how many permutations out of 20,000 gave as many or more hits than N'kisi actually obtained. For example, in the data shown in Table 2, N'kisi scored 23 hits. In a randomized permutation analysis, 23 or more hits occurred in only 5 out of 20,000 permutations. Thus, the p-value was $5/20{,}000 = 0.00025$.

2. In a Bootstrap Resampling Analysis (BRA) (Efron & Tibshirani, 1998), the probability of N'kisi saying a given word in a given trial was estimated by dividing the number of trials in which a key word was said by the total number of key words said (as in Table 3). Then, for each trial, the computer generated random "responses" in accordance with the probabilities that the various key words would be said. This random generation of responses for all trials was repeated 20,000 times, excluding repetitions, as in the RPA. The probability that N'kisi would have obtained the actual number of hits by chance was calculated from the number of "trials" out of 20,000 in which there were as many hits as those actually observed, or more. This BRA resembled the RPA procedure, but was much quicker, taking only minutes on the computer, rather than days.
3. It appeared that N'kisi repeated hit words more than miss words. To test this statistically, a 2×2 contingency table was constructed with "hits" and "misses" as the columns, and multiple responses and single responses as the rows. The single response figures were the totals of the words he said only once. The multiple response figures were the total number of times that N'kisi said key words when he repeated a key word during one and the same 2-minute trial. Thus, for example, taking the data as shown in Table 2, for multiple hits, in trial 4/2, "medicine" had a multiple response of 2, in trial 6/2, "hug," 3, and so on. This 2×2 table was analysed by means of Fisher's exact test (Siegel & Castellan, 1988).

Results

There was a remarkably good agreement between the three "blind" transcribers. An example of these transcripts for one of the trials is shown in Table 1. The picture Aimée was looking at in this trial was of a couple on a beach in skimpy swimwear.

Using a majority scoring method, when there was agreement between at least two out of three blind transcribers, in 71 out of 131 trials, N'kisi said one or more of the 19 key words (Table 2). In the remaining 60 trials, N'kisi either

TABLE 1
Independent Transcripts for Trial 25/3 By Anna Yamamoto, Betty Killa, Pam Smart, and Aimée Morgana[a]

Anna Yamamoto	Betty Killa	Pam Smart	Aimée Morgana
	Tones, whistles, creaks	Whistles and beeps	Beeps and whistles
Look at my pretty naked body	Look at my pretty *naked body*	Look at my pretty *naked body*	Look at my pretty *naked body*
	Whistle, creaks	Whistles and beeps	Beeps and whistles
Look at the little ... (?)	Look at the little ... Tones whistles, beeps	Whistles and beeps	Look at the little pict ... (-ure)
Look at my pretty naked butt	Look at my pretty *naked body*		Beeps and squeaks
	Tones, creaks, whistles		Beeps

Note: The key word "naked body" (scored as a single word) is present in all the transcripts (italic).
[a] Aimée Morgana's transcript is included for comparison.

remained entirely silent, or said none of the 19 key words corresponding to the test images. Thus, in these trials, neither a "hit" nor a "miss" was scored, and they were irrelevant to the analysis. Non-scorable comments made by N'kisi during these sessions were generally attempts to contact Aimée, or unrelated chatter about events of the day. Some of them, however, seemed to correspond to images Aimée was looking at during the trial, but such apparent "hits" could not be included in the statistical analysis because they did not involve prespecified key words.

In the 71 trials summarized in Table 2, N'kisi said 117 key words, of which 23 were hits.

In Table 3, the data are tabulated in rows corresponding to each of the 19 key words. The probability that N'kisi would have said a given key word by chance in any trial was calculated by dividing the number of times a given word was said by the total number of key words said, as shown in column E. The expected frequency of N'kisi scoring a hit by chance with this word was then worked out by multiplying the figures in column E by the number of trials involving a picture corresponding to the key word (column F). Adding up the expected frequencies of hits by chance in column F gave an estimate of the total number of hits expected by chance, namely 7.4. The actual number of hits was 23.

The data in Tables 2 and 3 were based on scores from a majority of the transcribers. When we considered only the scores where all three "blind" transcribers were in agreement, N'kisi said 105 key words, of which 19 were hits. When the scores included words recorded by only one out of the three transcribers, N'kisi said a total of 136 key words, of which 26 were hits.

In an RPA, all the key words N'kisi said were combined at random with the pictures used in each trial (as tabulated in Table 2). Because we did not count repetitions of key words in this analysis, random permutations that involved repetitions of a given key word were excluded.

For the data from the majority scoring method (Table 2), in 20,000 random permutations, there was a mean of 12.2 hits, with a standard deviation of 2.8. This was considerably lower than N'kisi's actual score of 23 hits. (This mean chance hit rate is more reliable than the chance hit rate of 7.4 shown in Table 3, arrived at by adding up the estimated hit frequencies for the different words.)

In this RPA, there were 23 or more hits in only 5 out of 20,000 permutations. Thus the probability of N'kisi scoring 23 or more hits by chance was 5/20,000, or 0.00025 (Table 4, row I.B, RPA column). For comparison, in Table 4 we also show the probabilities using the scores when all three scorers were in agreement (Table 4, row I.A), and including key words detected by only one of the scorers (Table 4, row I.C). In all cases, the results were highly significant statistically.

A Bootstrap Resampling Analysis gave very similar results to the RPA (Table 4, I). For example, with the data from the majority scoring method (Table 2), out of 20,000 random permutations, there was a mean of 11.8 hits, with a standard deviation of 2.9. There were 4 out of 20,000 permutations with as many hits as N'kisi or more, giving a p-value of $4/20{,}000 = 0.0002$ (Table 4, Row I.B. BRA column).

The list of N'kisi's vocabulary from which key words had been chosen was not edited for frequency or reliability of use, and included some words that N'kisi had used only rarely, and did not utter at all during this series of trials. These words were "cards," "CD," "computer," "fire," "keys," "teeth," and "TV." There were 18 trials involving pictures corresponding to these words in which N'kisi could not have scored either a hit or a miss, since he never said these words. In established practices for testing language-using animals, the words tested are typically screened in some way for reliability of production (Fouts, 1997). Perhaps a better way of analyzing the results would be to exclude the 18 trials involving these images. The results of this analysis are shown in Table 4, II. This method reduced the number of misses, and consequently the proportion of N'kisi's hits increased. For example, by the majority scoring method (B), 23 words out of 82 were hits (28%). Nevertheless, this method made little difference to the statistical significance of the results, as shown by a comparison of parts I and II of Table 4.

One of the reviewers of this paper pointed out that much of N'kisi's success hinged on the frequency of his hits with the word "flower," which was both the most common key word he said during the series of tests, and was also represented by the largest number of images. If this word were to be excluded from the analysis, the statistical significance of N'kisi's success would be lower. This is true, but post hoc. In any experiment, if the most obvious evidence is arbitrarily removed afterwards, the significance of the results will be reduced. Nevertheless, to examine this argument more closely, we carried out a statistical analysis eliminating "flower" both as a target and as a response. Using the data from the majority scoring method, as shown in Table 2, following the BRA procedure with 20,000 random permutations, the results excluding "flower" were still strikingly significant ($p = 0.006$).

TABLE 2
Data for Trials in which N'kisi Said at Least One Key Word, as Scored by a Majority of the Independent Transcribers

Trial	Picture	Key words said[b]
4/1[a]	CD	hug, medicine, doctor (2), water
4/2	medicine, bottle, doctor, glasses	*medicine (2)*
5/1[a]	teeth	water (2), doctor (2)
5/2[a]	CD	doctor
5/3	bottle	car (2), doctor (2)
5/4	water	doctor (2)
5/5[a]	CD	car, doctor
6/2	*hug, water*	*hug (3), water*
6/4[a]	keys	hug (3)
8/1[a]	CD	medicine
8/2	hug, glasses	*glasses,* medicine (2), doctor
8/3	bottle, medicine	*medicine* (2), doctor (2)
8/4[a]	CD	glasses, car, doctor (2)
8/5	fire, glasses	hug, medicine
9/1[a]	fire	medicine
11/3	flowers	water, medicine
11/4	book	water, phone
12/1[a]	computer	water, medicine (2)
12/3[a]	CD	phone
12/4	phone	*phone* (2), medicine, water
13/2	car	*car*
15/3[a]	computer	phone, car, naked body
15/4	water	feathers
15/5	flowers	car, water
16/1	feather	car
16/2[a]	key	medicine, glasses
16/3	flowers	glasses (3)
18/1[a]	cards	water
18/2	water, hug, naked body	*water*, glasses
18/4	phone	glasses
19/1	car	medicine
19/3	phone	doctor
19/4	flower	*flower,* medicine
19/5	car	flower
20/3	books	doctor
21/1[a]	fire	flower (4)
21/2	flower	*flower* (10), doctor (2)
21/3	flower	*flower (7)*
21/4[a]	computer	flower (4), books
21/5	flowers hug	*flower (4)*
22/2[a]	TV	glasses
23/1	flower	*flower,* doctor, glasses (2)
23/2	flower	*flower,* car
23/3	water	flower
23/4[a]	key	glasses, feather
24/2	flower	naked body
24/3	naked body, water	*naked body*
24/4	doctor, medicine, glasses	*medicine,* feather, naked body, water
24/5	books	glasses
25/3	naked body, water	*naked body*
25/5	naked body, water	*naked body (4)*
27/1	flower	*flower*

TABLE 2
Continued

Trial	Picture	Key words said[b]
27/2[a]	CD	flower, doctor (3), naked body
27/3	hug	medicine, doctor (2)
28/1	car, book	*book,* flower
28/2	bottle, medicine	flower (4)
28/3	hug, water, glasses	flower (2)
28/4	flower	*flower,* naked body
28/5	flower	*flower* (3), bottle
29/1	flower	water
29/3	phone	flower, naked body, feathers
30/2	flower	*flower,* bottle
30/3	phone, glasses	flower, bottle
30/4	flowers	naked body, doctor
31/1	bottles	book
31/3	water	flower, medicine
31/4	glasses	naked body
31/5[a]	computer	hug, flower
33/1	phone	flower (3)
33/3	hug, flowers	phone
33/5	water	phone (2)
hits	excluding repetitions	23
misses	excluding repetitions	94
hits	including repetitions	51
misses	including repetitions	126

[a] The trials marked by an asterisk involved pictures corresponding to key words that N'kisi rarely used, and never said in the entire series of trials.
[b] Some "key words said" were repeated during a given trial, and the number of times N'kisi said the key word are shown in parentheses; hits are italicized.

Repetitions

In the analyses described above, we ignored the repetitions of key words during a given trial. But looking more closely at N'kisi's responses, we noticed that he tended to repeat hit key words more than he repeated misses. Although this is a post hoc analysis, it seemed worthwhile to examine the repetitions in more detail to obtain more information about N'kisi's use of language during the tests.

In the data from the majority scoring method (Table 2), N'kisi repeated the key words in 9 out of 23 hits (39%), and in 22 out of 94 misses (23%). Also, the number of multiple responses per word was greater with repeated hits than repeated misses: an average of 4.1 as opposed to 2.5.

The total numbers of multiple and single responses with hits and misses are shown in Table 5. N'kisi repeated hits very significantly more than misses ($p = 0.0003$ by Fisher's exact test).

Discussion

By all scoring methods and by all methods of statistical analysis, N'kisi scored very significantly more hits than would have been expected by chance.

TABLE 3
Results of the N'kisi Telepathy Trials Based on Key Words, as Scored by a Majority of the Independent Transcribers[a]

	A. image	B. no. of images	C. no. of times word said	D. no. of hits	E. p word = C/117[b]	F. ef hit = E × B[c]
1	Book	4	3	1	0.026	0.103
2	Bottle	5	3	0	0.026	0.128
3	Car	4	8	1	0.068	0.273
4	Cards	1	0	0	–	–
5	CD	7	0	0	–	–
6	Computer	4	0	0	–	–
7	Doctor	2	16	0	0.137	0.274
8	Feather	1	4	0	0.034	0.034
9	Fire	3	0	0	–	–
10	Flower	17	23	10	0.197	3.341
11	Glasses	7	10	1	0.085	0.598
12	Hug	7	5	1	0.043	0.299
13	Keys	3	0	0	–	–
14	Medicine	4	16	3	0.137	0.547
15	Naked body	4	11	3	0.094	0.376
16	Phone	6	6	1	0.051	0.308
17	Teeth	1	0	0	–	–
18	TV	1	0	0	–	–
19	Water	10	12	2	0.103	1.128
	Totals	90	117	23	–	7.409

[a] This table includes data from all trials in which N'kisi said at least one key word, as shown in Table 2, ignoring repetitions.
[b] Column E shows the probability of N'kisi saying a given key word by chance on any particular occasion, calculated by dividing the number of trials in which the word was said (column C) by the total number of words said (117).
[c] Column F shows the expected frequency (ef) of N'kisi achieving a hit by chance, calculated by multiplying the probabilities in column E by the number of images, as shown in column B.

TABLE 4
Statistical Significance of N'kisi's Hits in the Telepathy Experiment, Using Data from the Three Independent Transcribers[a]

Transcribers	Hits	Misses	p (RPA)[b]	p (BRA)[b]
I. Including all 19 key words				
A. 3	19	86	0.004	0.004
B. 2	23	94	0.0003	0.0002
C. 1	26	110	0.0002	0.0002
II. Including only the 12 key words actually said in one or more trials				
A. 3	19	55	0.005	0.004
B. 2	23	59	0.0005	0.0003
C. 1	26	73	0.0003	0.0003

[a] For the data in row A, all three transcribers were in agreement; in row B, at least two were in agreement; and row C includes key words that were detected by only one out of three.
[b] In the Randomized Permutation Analysis (RPA) and Bootstrap Resampling Analysis (BRA), the probability of the observed number of hits arising by chance was calculated on the basis of how many times out of 20, 000 there were as many hits as N'kisi actually obtained.

TABLE 5
Multiple and Single Responses by N'kisi with Hits and Misses[a]

Responses	Hits	Misses	Total
Multiple	37	54	91
Single	14	72	86
Total	51	126	177

[a] Data from the majority scoring method, as in Table 2.

As far as we know, testing a language-using animal for telepathy has never been attempted before. Clearly, an animal cannot be expected to understand and respond to a testing situation exactly as a human would.

For at least two reasons, our test procedure probably underestimated N'kisi's hits. First, we could not explain to him that he was being tested in a series of two-minute trials, and that when the two minutes was over he should stop saying some words and start saying others. On 13 occasions, either he went on saying a word that had been a hit after the following trial began, or else he said a word corresponding to the previous photograph in the following trial. We counted these words as misses, but they may have been repetitions of hits, or delayed hits.

Second, N'kisi could not have understood the need to limit his responses to the list of key words we had specified in advance. He has never been trained to produce specific words on demand; instead, Aimée had always allowed him to use language as he pleased. In at least 10 trials, he said words or phrases that corresponded to pictures Aimée was looking at but which did not involve prespecified key words. For example, in one trial Aimée was looking at a photograph of a stationary car whose driver had his head out of the window. This photograph had been chosen to represent the key word "car." N'kisi did not say "car." Instead he said, "Uh-oh, careful, you put your head out," the moment Aimée noticed this unusual detail. This was not counted as a hit (or as a miss) because the phrase "head out" was not on the list of prespecified key words.

But even though our procedures probably underestimated N'kisi's performance, the results were highly significant statistically and imply that N'kisi was influenced by Aimée's mental activity while she was looking at particular pictures, even though he could not see her, hear her, or receive other "normal" sensory clues. N'kisi's very significant repetition of key words when they were hits adds to this conclusion.

The statistical significance of our results incidentally confirms N'kisi's meaningful use of spoken language. If N'kisi were incapable of using language appropriately, it would probably not have been possible to come up with any significant results in this series of tests. The ability of animals to use language, and what this may reveal about their cognitive abilities, is still being debated. These results raise issues relating to animal intelligence and interspecies communication with implications that extend beyond the scope of this paper.

These experimental tests are consistent with Aimée's frequent observations of seemingly telepathic behavior by N'kisi in everyday situations. Aimée has also noticed this with other parrots she has kept, and other parrot owners have found that their birds sometimes seem to respond telepathically to their thoughts and intentions (Sheldrake, 1999, 2003).

Several other animal species kept as pets show behavior that seems telepathic. For example, some dogs and cats anticipate the arrivals of their owners by going to wait at a door or window 10 minutes or more before the person arrives home, and exhibit this anticipatory behavior when the person is still several miles away. In a series of videotaped experiments with dogs, in which the owners came home at randomly chosen times, it was possible to rule out explanations in terms of routine or sensory clues. The dogs still anticipated their owners' arrivals when they came home at non-routine times, and when no one at home knew when to expect them (Sheldrake, 1999; Sheldrake & Smart, 1998, 2000a,b). Sensory clues were ruled out in tests in which the owner returned home at non-routine times in unfamiliar vehicles, such as taxis, from places more than 5 miles away. The dogs appeared to be reacting to their owners' intentions (Sheldrake & Smart, 1998, 2000a,b).

There is great potential for further research on telepathy in animals. Language-using parrots have the advantage of being able to speak, and, as a result, are capable of a greater diversity of responses than dogs, cats, and other domesticated species. We are continuing our research with N'kisi.

Acknowledgments

We thank Betty Killa, Pam Smart, and Anna Yamamoto for transcribing the tapes, and Evan Izer for organizing the photo source materials. We are very grateful to Jan van Bolhuis for carrying out the statistical analyses, and for his helpful advice. This work was supported by the Bial Foundation, Portugal; the Lifebridge Foundation, New York; and the Institute of Noetic Sciences, California.

References

Efron, B., & Tibshirani, R. J. (1998). *An Introduction to the Bootstrap.* Boca Raton: Chapman and Hall.

Fouts, R. (1997). *Next of Kin: My Conversations with Chimpanzees.* New York: Avon.

Patterson, F., & Lindin, E. (1981). *The Education of Koko.* New York: Holt, Rinehart and Winston.

Pepperberg, I. (1999). *The Alex Studies: Cognitive and Communicative Abilities of Grey Parrots.* Cambridge, MA: Harvard University Press.

Sheldrake, R. (1999). *Dogs that Know When Their Owners are Coming Home, and Other Unexplained Powers of Animals.* New York: Crown.

Sheldrake, R. (2003). *The Sense of Being Stared At, and Other Aspects of the Extended Mind.* New York: Crown.

Sheldrake, R., & Smart, P. (1998). A dog that seems to know when his owner is returning: Preliminary investigations. *Journal of the Society for Psychical Research, 62,* 220–232.

Sheldrake, R., & Smart, P. (2000a). A dog that seems to know when his owner is coming home: Videotaped experiments and observations. *Journal of Scientific Exploration, 14,* 233–55.
Sheldrake, R., & Smart, P. (2000b). Testing a return-anticipating dog, Kane. *Anthrozoos, 13,* 203–212.
Siegel, S., & Castellan, N. J. (1988). *Nonparametric Statistics for the Behavioral Sciences.* New York: McGraw-Hill.

Reviewer Comment: Jeffrey Scargle

In reviewing this paper, I commented on the preponderance of flowers in two aspects of the experiment. The parrot's vocabulary and the selection of images are not random. Both show enhanced frequencies of certain culturally favored concepts—"flower" being only the most obvious one. The statistical analysis presented in the paper assumes randomness, and is therefore inappropriate.

The authors' response (that it would be unfair, statistically speaking, to arbitrarily remove the most obvious evidence) is correct but misses the point. For their comment to be relevant, "flower" and other concepts, such as "doctor," and "medicine," would have to be the most frequent concepts in this experiment simply because of statistical fluctuations. But both samples (the parrot vocabulary and suite of pictures) contain them with higher than average frequency because of their cultural importance.

It is very difficult, if not impossible, to account for this kind of cultural selection effect with statistical analysis of the data. Hence, I do not believe that this experiment provides any evidence supporting the claim of telepathy. What is needed is an experimental protocol insensitive to selection effects in the first place.

Analyzing the data kindly provided by the authors, I found a p-value 4 times larger than theirs for the case where "flower" is removed, but as discussed above this calculation is irrelevant.

Jeffrey D. Scargle
Space Science Division
NASA Ames Research Center
e-mail: jeffrey@cosmic.arc.nasa.gov

Reviewer Comment: Mikel Aickin

Sheldrake and Morgana have done an admirable job in an area of research that is plagued with both methodological and statistical problems. As is customary in *JSE*, they have described their procedures in meticulous detail, a salutary strategy that becomes particularly important when the compulsive skeptics start sniffing around.

When I originally refereed this article, I was concerned mainly by the omission of the instances in which N'kisi said nothing. It seemed to me that opportunities for him to have had a match, but where he failed, should be

counted as failures, regardless of whether he said anything or not. I therefore requested data on the omitted cards/phrases, which the authors immediately supplied. I did a permutation test on the entire dataset, and found a p-value that differed only trivially from the one stated in the article. Although the authors have done an analysis that I would not have done (by omitting data), it makes no difference to the results, and so I was happy.

In passing, I mention that the permutation test done in the article is incorrect. As described, it permutes all 117 words and then assigns them to images with the same frequency as was actually observed (thus, the first image was assigned to four words). This procedure makes the null hypothesis a combination of two separate hypotheses: (1) that N'kisi responds at random, and (2) that the words he uses in any particular session are selected at random. It is only (1) that is at issue, and adding in (2) is a methodological error. The error is that N'kisi could combine words dependently, causing rejection of the hypothesis (1) and (2), leading the researchers to reject (1), when it is (2) that causes the rejection. I make this point only for methodological reasons, because as I said above, a proper permutation test does not contradict the author's conclusions.

I am not bothered by Scargle's concerns. The fact that some words occur more or less frequently in the cards or in the responses is irrelevant (so long as it doesn't become ridiculous, like having a flower on every card). The permutation test (in its correct manifestation) remains appropriate in this case. Sheldrake and Morgana have demonstrated, in a particular instance, a concordance between N'kisi's phrases and the card images Morgana viewed cannot be explained by chance and does not appear to be explainable by methodological error.

Mikel Aickin
e-mail: Mikel.Aickin@kpchr.org

Editorial Comment

This article is another instance of your Editor's difficulties where research protocols and statistical inference are questioned. The first two reviewers of this manuscript made opposing recommendations, and two more reviewers were consulted. Publication was a majority recommendation, though some reviewers felt that the protocol was flawed for the reason described above by Scargle. Aickin describes how he grappled with the issues and was helped by further information from the authors.

My personal reaction is that, once again, we have suggestive results, a level of statistical significance that is less than compelling, and the devout wish that further work with refined protocols will ensue.

At the same time, I want to express publicly our deep indebtedness to reviewers who have, time and again, spent much time and effort in clarifying issues, stimulating authors to refine their presentations, and informing readers of the various views that can be legitimately taken on some of these matters.