INTERACTIVE HUMAN COMMUNICATION

Studies of how people communicate in solving problems assist progress toward the development of a conversational computer, with which the user could communicate as if it were a person

by Alphonse Chapanis

Modern computers touch the life of every citizen in varied and often unexpected ways. Not only do computers prepare our utility bills, credit-card bills and bank statements but also they control our traffic, assist us in making travel and theater reservations, keep tabs on the weather for us and help to diagnose our bodily ills. For all that, most of us still have little direct contact with computers. Most computers still require an intermediary between the ultimate user and the computer, someone who is familiar with the way the computer works and with the special language that is needed to address it.

A goal toward which many people have been working is the design and construction of conversational computers: computers that can interact with people in such familiar and humanlike ways that they require little or no special instruction. If such conversational computers are ever to come into existence, however, their designers and programmers will need to know more about how people interact in communicating with each other. With this rationale in mind my colleagues and I at Johns Hopkins University have been working to describe human communication in precise terms and to define its rules.

We have been concerned with three main questions. How do people naturally communicate with each other when they exchange factual information in the solution of problems? How is interactive human communication affected by the devices through which people converse? What other significant variables affect interactive communication?

Let me digress briefly to distinguish between unidirectional and interactive communication. For years psychologists have been concerned with the effectiveness of unidirectional modes of communication such as highway signs, books, lectures and television programs. In unidirectional communication the person to whom the message is addressed is a passive recipient of information. Nothing that he says or does affects the communicator, the communication process or the content of the message.

Interactive communication involves at least two participants. The content of any particular message is determined in part by the content of the prior messages from all participants and so cannot be predicted from the content of the message from any one of them. Conferences, arguments, seminars and telephone conversations are examples of interactive communication. This is the kind of communication that has been the focus of our investigation.

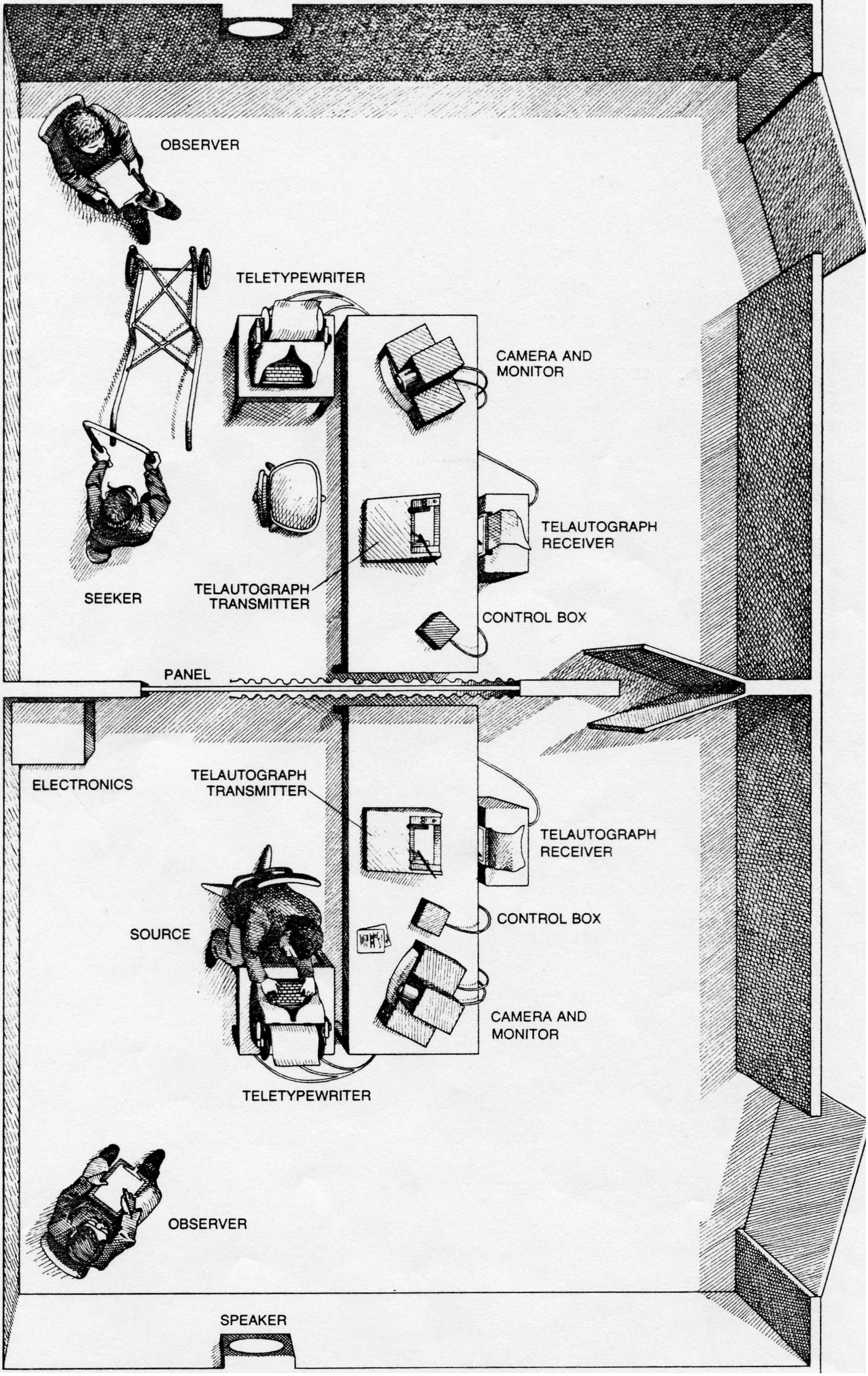
Our experiments are designed to model interactions between man and computer rather than to simulate any existing or planned interactive computer systems. We set up two-person teams and ask them to solve credible problems for which computer assistance has been or could be useful. The exchanges that result represent a limited class of conversations, to be sure, but it is an important class, and we have to start somewhere.

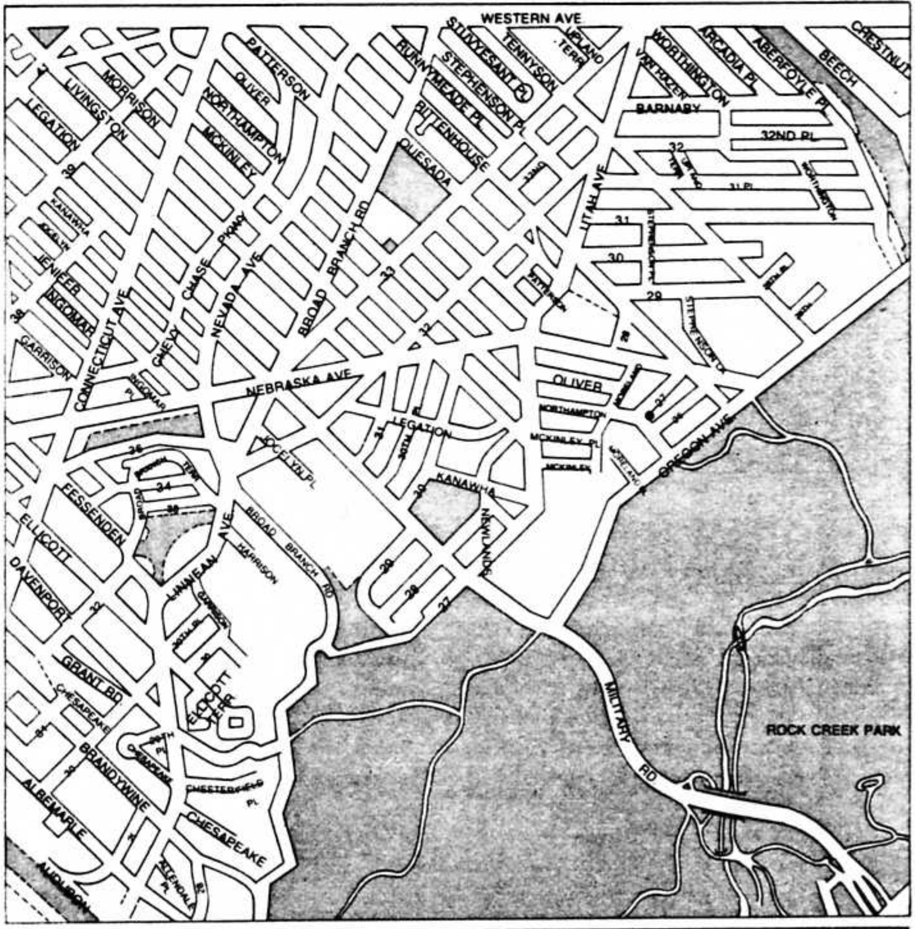
One of our primary interests is the channels and the modes through which people converse. Although the channels of communication that link man and computer are being broadened, most interactions of this kind involve a typewriter or a similar device. Our experiments examine four different channels: voice, handwriting, typewriting and video, the last being the picture part of television without the voice. Three of these four basic channels have been tested singly, and all of them have been tested in various communicative combinations that we call modes. We have tested as many as 10 different modes in a single experiment. As a standard of comparison we typically rely on normal, unrestricted, face-to-face communication, which we call a communication-rich mode.

When we have set up a team, we designate one member as the source (of information) and the other as the seeker. One can think of the source as an ideal computer, that is, a computer communicating in such a human way that a person who did not know he was dealing with a computer might readily believe he was communicating with another person. The seeker can be regarded as the user of the computer. To continue the analogy, our different channels and modes of communication model various input and output channels between the computer and its human user.

The setting for a typical experiment consists of two adjoining rooms connected by a soundproof double door [*see illustration on opposite page*]. The wall between the rooms also has in it a large double-glass panel, which can be covered with an opaque screen so that the source and the seeker cannot see each other. When the panel is not covered, the participants can see each other and can converse freely through a microphone and loudspeaker, but they are still separated physically. Some of our experiments also have test conditions in

LABORATORY SETTING for a typical experiment is depicted on the opposite page. The seeker has been given a trash-can carrier to assemble but has not been told its name or function. The source has the information for assembly. The experiment is designed to elicit communication in the hope of assisting in the design of a computer that would be analogous to the source in communicating much as a person communicates.


OBSERVER
TELETYPEWRITER
CAMERA AND
MONITOR
TELAUTOGRAPH
RECEIVER
SEEKER
TELAUTOGRAPH
TRANSMITTER
CONTROL BOX
PANEL
ELECTRONICS
TELAUTOGRAPH
TRANSMITTER
TELAUTOGRAPH
RECEIVER
CONTROL BOX
SOURCE
CAMERA AND
MONITOR
TELETYPEWRITER
OBSERVER
SPEAKER



Earich La. I-25	Milmarson Pl. R-11	Tuckerman St...Q-9, R-10, R-12	41st St., S.E. O-24
Earl Pl., N.E. R-17	Mills Ave., N.E. Q-17, R-17	Tulip St. R-6	42nd Pl. M-6
East Pl. K-13	Millwood Lane J-8	Tunlaw Road K-10, 11	42nd Pl., N.E. R-21
East Capitol St. M-18, Q-22	Milwaukee Pl., S.E. J-22	U Pl., S.E. M-22	42nd Pl., S.E. O-24
East Capitol St., N.E. O-21	Minnesota Ave., N.E. Q-21, R-21	U St. I-9, K-11, M-13, N-15	42nd St., N.E. R-21, Q-22
East Beach Drive R-7	Minnesota Ave., S.E. M-22, O-22	U St., S.E. L-21, N-23	42nd St. M-7, L-8, J-10
Eastern Ave. S-7, S-10	Mintwood Pl. M-12	U St., S.W. J-19	42nd St., S.E. O-24
Eastern Ave., N.E. S-21		U St., N.E. O-16	43rd Pl. L-7
Easy Pl., S.E. P-24			43rd Pl., N.E. R-21

Allin John O 3800ResrvrRdNw ---------625-7351
If no answer call ----------------223-2200
Alper Melvin G 2141 K Nw ------------223-3300
If no answer call ----------------296-4516
Alpert Hubert J MD 1234 19thNw ------223-5560
If no answer call ----------------223-2246
Res 4761BradlyBlvd ChCh ----------652-4493
Angeles Ramon Jr MD—
Hours By Appointment Only
461 H Nw -------------------- 737-4232
If no answer call Medical Bureau ----223-2200
Angelos Peter G Dr

ORIENTATION PROBLEM imposes on the seeker the task of finding the address of the physician closest to seeker's hypothetical home. Seeker is given a street map of Washington, D.C., with the address marked as shown by black dot. He also receives a street index (*middle*) keyed to the map. The source receives one page from the listing of physicians in the yellow pages of the Washington telephone directory (*bottom*). Subjects occupy separate rooms and must solve the problem by one or more modes of communication.

which the two people can neither see each other nor communicate by voice; instead they use writing machines linked in such a way that anything typed or written in longhand on one machine is reproduced on the other.

Our problem-solving tasks differ significantly from the kind usually found in the problem-solving literature of psychology because they were designed to meet certain special criteria. They sample such psychological functions as verbal skill and psychomotor skill. They are representative of tasks for which interactive computer systems are or could sometimes be employed. Instead of being abstract or artificial puzzles of the kind often devised to measure hypothetical psychological processes, they are of recognizable and practical importance in everyday life. They have definite, recognizable solutions, which can usually be reached within approximately an hour. Finally, their solution requires no special skills or specialized knowledge.

The tasks are formulated in such a way that solving them requires the seeker and the source to work together as a team. The seeker is given a problem for which he has to find the solution. His information folio consists of certain parts of the problem. The source has a folio with the remainder of the information needed to solve the problem. Neither person can solve the problem by himself, but together they have all the information needed for doing so. Remember, however, that our problems are designed to elicit communication between the two members of a team. They do not necessarily represent the way tasks would be assigned to man and computer in any real system.

All together we have constructed 10 problems that meet our needs. The following brief descriptions of three of them will convey their flavor.

In the "equipment-assembly problem" the task of the seeker is to assemble a common household article: a trash-can carrier. His information folio consists of all the disassembled parts of that article exactly as it comes from the mail-order house from which it was bought. He is not told either the name or the function of the device. The source's folio consists of the set of diagrams and instructions for assembly that came with the parts.

In the "information-retrieval problem" the seeker has to find the citation of every newspaper article relevant to an assigned topic that appeared in *The New York Times* during a given year. Usually he is told not to count editorials, reports of public speeches or letters to the editor. The source's information folio consists of *The New York Times Index* for the same year.

In the "geographic-orientation problem" the seeker's task is to find the office or residence address of the physician closest to a hypothetical home address. He is supplied with an index of streets, a gridded street map of Washington, D.C., and a card on which the home address is typed. His hypothetical home address is also marked on the map. The source is supplied with one page of the list of physicians in the yellow pages of the Washington telephone directory.

Our subjects have varied from experiment to experiment. We have relied heavily on that mainstay of psychological experiments, the college student. In one experiment, however, we enlisted high school boys, in another girls from a parochial high school and in a third a mixture of college and high school students.

We began with a series of tests involving four types of communication with less sophisticated equipment and procedures than have characterized our later experiments. In the communication-rich mode the subjects sat side by side at a table with no barrier between them. In the voice mode they were in separate rooms and communicated through a cloth panel that could be heard through but not seen through. In the handwriting mode they wrote messages in a notebook they passed through a slot in the wall between the two rooms. In the typewriter mode we had both experienced and inexperienced typists.

The results show large differences among the several modes of communication [*see illustration on next page*]. The inexperienced typists, for example, took almost two and a half times as long to solve problems as subjects in the communication-rich mode did. Differences of the same order have turned up repeatedly in other experiments.

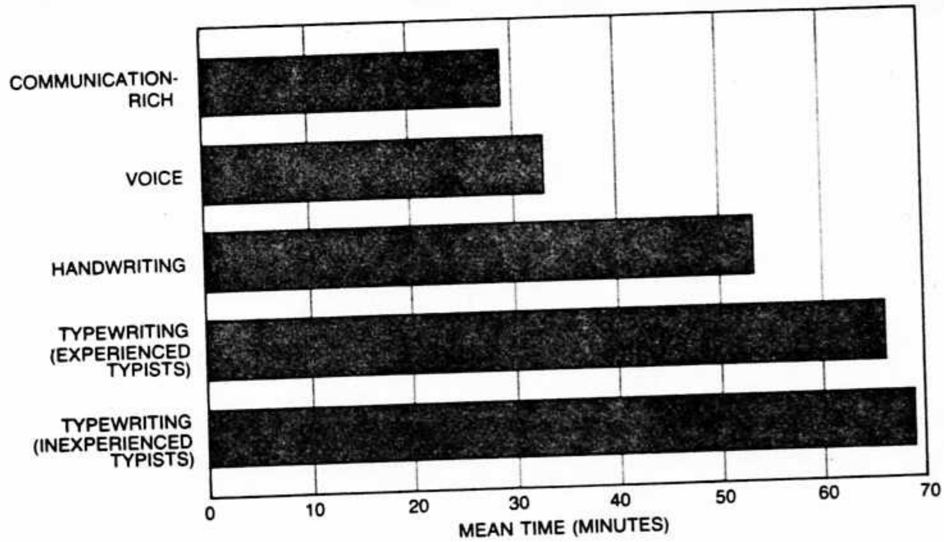
An unexpected finding was the notably small difference in performance between the experienced and the inexperienced typists. This finding seemed so implausible that we later checked it

```
goaheaddoyouknowhowto put this togher
ill tryits a trash toter  ill type you the directions ok
put axle thru 38th holes from outside
38th holes/ ??yes
put 1 handlebar on back of  each outer frame line up bolt holes
what does outer frame look like ? its like a (W)
put bottom frame to outer frames on front + rear of outer frames
ok use 1+12 bolts
are your parts labled by lettrs ???
nookthe thing looks like a cart with room for 2 trash cans the part
that looks like this(XX)goes on the bottom +the 2(W)parts go on the sids
put male ends ? into female ends
what does that mean? i dont no
it looks like 3(u)s
what? 2(u)s go into each other then theyare put on other u +put
on W put top frame to front of outer fr.+to handlbar 2 1/4
bolts put center support fr. inside topfr. use 2 1/4 bo. thru
center of top fr. put 2 1/12 bolts thrub center of side fr.,
bottomfr. 2 bottom of center support fr.
okput on wheels 3 spoks on outside put on hubcap with hammer
put oh handgrips  DO ALL THESE STEPS FOR BOTH SIDES   ok?????
```

TYPED EXCHANGE between a seeker and a source engaged in solving the equipment-assembly problem is reproduced in part, with messages by the seeker underlined. In spite of such confusing messages as "38th holes" in the third line, where the source intended to say 3/8th-inch holes, the two members of the team solved the problem in less than an hour. Both were inexperienced at typing.

```
  SO:  Okay.  And but it's, all it is, is a frame.
  SK:  All it is, is a frame.  And what's supposed to fit inside the frame?
       Do you fit, do you put, you know, a pan like in a wheel barrel or/
  SO:  No, there's no pan.
{ SO:  Okay, now.
{ SK:  Or is it [Message continues below.]
  SK:  like a wagon or what?
  SO:  Now let, now let me read this for a second..It's like a wheel bar-
       rel like I said, there's a handle, there's two wheels..and all it
       is is like it's a frame wheel barrel like. But there is no, y'know,
       water will go through it in other words.
  SK:  Do wheels go in the front?
  SO:  I'm gonna, I'm gonna [Message continues below.]
{ SO:  read this.
{ SK:  Or do the [Message continues below.]
  SK:  wheels go in the back?
  SO:  I'm gonna read the direction now.  I'm gonna, give you, the, you
       know..Let me read it..oh, trash toter. Oh, is that what it is?
```

VOICE EXCHANGES between a seeker and a source working on the equipment-assembly problem are reproduced in a partial transcript in which *SO* is the source and *SK* the seeker. A bracket linking *SO* and *SK* indicates that both were talking simultaneously.



SOLVING TIME of problems is averaged for several modes of communication. In the communication-rich mode, for example, where the two members of a team were in the same room and could communicate freely, the average time for a solution was 29 minutes.

in another experiment with different subjects and more elaborate procedures to help us figure out what was going on. We now think the explanation has at least two different components.

First, by means of detailed measurements of what the subjects were actually doing when they solved problems we found that the average subject spends somewhat less than a third of his time in communication. In interactive problem solving subjects do a great many other things, such as make notes, think about what to say, handle objects and search for information in their respective folios. As a result the advantages one might expect to come from superior typing skill are diluted by the relatively small fraction of time in which the skill can be exercised.

Second, the kind of typing called for in interactive communication is unusual. Typing skill is normally measured by having subjects copy text material. In communication by typewriter, however, the communicator has to decide what to say, compose his thoughts into a message (often fragmented and incomplete) and then type out the message. The transmissions are characterized by hesitations, mistakes, changes of thought and irregular tempos that may at various times be indirect expressions of doubt, amusement or anger. In short, typing skill is usually measured as a strictly mechanical or psychomotor activity, whereas communication by teletypewriter is a much more intellectual process. It is small wonder that the two techniques seem to have so little in common.

From the voluminous literature on kinesics, gestures and "body language" I had been led to predict a large difference between face-to-face communication and communication by voice alone. The voice channel by itself seems impoverished in comparison with the variety and richness of the information-bearing clues available in face-to-face communication. The data did not conform at all to my expectations. The average amount of time taken to solve problems by voice alone was only slightly more than it was in face-to-face communication.

If this were an isolated result, one might well question its validity. It has appeared, however, in other experiments that we have done with different subjects and different problems. One of these experiments, carried out by a former student of mine, Robert B. Ochsman, tested 10 different modes of communication. Arranging them according to richness, one finds that here too the mean problem-solving time is only slightly longer in the voice mode than in the communication-rich mode [*see illustration on opposite page*].

Although solution times tend to increase as the channels of communication become more impoverished, the most striking feature of our data on the 10 modes is that they tend to fall into two fairly distinct groups. The faster five all have a voice link, whereas the slower ones do not. Statistical tests confirm that this one comparison is the only statistically significant one among the 10 communication modes.

However interesting the data on problem-solving time are by themselves, they become even more interesting when they are related to the linguistic output of the communicators. The problems of measuring and quantifying that output have in turn been almost as interesting as the results. Most psycholinguistic research has been done on what I call immaculate prose. Such prose consists of grammatically correct sentences with nouns, verbs and other parts of speech in their proper place. Words are spelled correctly and rules of punctuation are observed. All computer programs based on what is called natural language require immaculate prose because the sentences that are fed into the computer are parsed in one way or another so that the meaning of the ensemble can be inferred from conventional rules of syntax.

The trouble is that people do not naturally speak in sentences. Most of us realize this in an intuitive way, but I suspect that few of us appreciate just how untidy normal human conversations really are. In our experiments recordings are made of conversations in all communication modes having a voice channel, and the recordings are then transcribed into typewritten protocols. Subjects generate their own protocols when they converse in the handwriting and typing modes.

Looking at the transcripts of conversations by voice or writing, one readily sees the untidiness I have mentioned. An example is the record produced by two inexperienced typists who were solving the equipment-assembly problem [*see top illustration on preceding page*]. The first impression is one of complete unruliness. Not one grammatically correct sentence appears in the entire protocol. Words are continually misspelled and run together. Abbreviations, both conventional and unconventional, are common, and violation of the rules of punctuation is commoner than their observance.

The record even includes serious errors of fact. For example, at one point the source gives the seeker an instruction about "38th holes." The seeker queries the source on this point, and the source replies that his original statement was correct. Actually what the assembly instructions stated and the source meant to say was that the seeker should put the axle through the 3/8th-inch holes.

Perhaps the most remarkable thing is that in spite of all this apparent unruliness, the information got through. The two team members who generated the protocol completed their task successfully in less than the average time required by teams for the equipment-assembly problem. Equally remarkable is that the protocols of the experienced typists were also characterized by many of the same kinds of error and ungrammatical feature. Evidently unruliness tends to creep in when the emphasis is on precision of typing.

Transcripts of voice conversations have their own special idiosyncrasies that are no less perplexing and difficult to deal with [*see bottom illustration on page 39*]. It is clear that if truly interactive computer systems are ever to be created, they will somehow have to cope with the mispronunciations, errors and violations of format that are the rule rather than the exception in normal human communication. Discovering the rules and characteristics of normal communication is a problem that has been ignored by linguists for too long.

Measuring and counting such characteristics of the protocols as words, sentences and messages had seemed simple in prospect but proved difficult in execution. In the end, however, we were able to formulate sets of rules that enabled us to count the linguistic units. Next came the task of trying to decide what measures of linguistic performance to apply to our linguistic units.

On the basis of hunches, hypotheses and what we could find in the psychological literature we came up with 136 linguistic measures. A number of them turned out to be trivial, and many others were so highly intercorrelated that they were redundant. In the end we were left with only nine meaningful measures of linguistic performance that describe our data. To a certain extent the discarded measures are as interesting as the useful ones because they reveal where it would be fruitless to expend time and effort in the future. To list them here, however, would consume too much space.

The productive measures are: (1 and 2) The number of messages generated by each subject and, closely correlated, the number of sentences. (3 and 4) The number of words per message and, closely correlated, the number of words per sentence. (5) The percentage of sentences that were questions. (6) The total number of words employed by a subject. (7) The total number of different words employed by a subject. (8) The ratio of different words to total words, called the type-token ratio. (9) The communication rate, which is the number of words communicated per minute of time actually spent communicating.

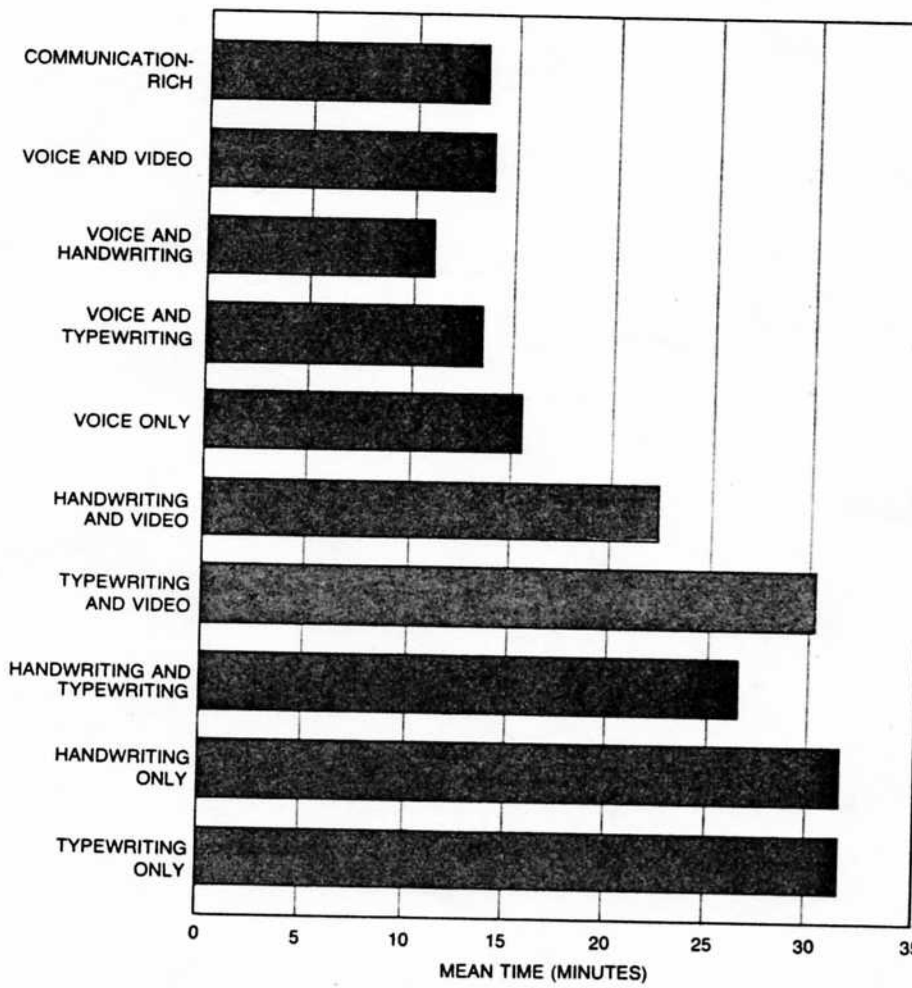
In one sense our findings are disappointing, since there appears to be so little to show for so much effort. In another sense, however, they are gratifying. The linguistic performance of people who communicate naturally can be described by a rather small number of quantitative measures.

When the data are summarized [*see illustration on next page*], the most striking thing about them is that the two fast modes of communication (the two that have a voice channel) are also extremely wordy. Subjects using the two voice modes, as compared with handwriting and typing, delivered about eight times as many messages, eight times as many sentences, five times as many words and twice as many different words; they also communicated words at a rate nearly 10 times as fast. In sum, the voice modes are fast ways of communicating, but they are extremely wordy, no matter how wordiness is measured.

The higher type-token ratio for the handwriting and typing modes confirms that they tend to be more parsimonious and less redundant than the voice modes. (Measuring redundancy in these experiments is extremely difficult because the data do not lend themselves to the conventional measures of redundancy relied on by information theorists. By making certain plausible assumptions, however, we concluded that in the voice modes subjects employ about 13 times as many words and four times as many unique words as are really required to solve problems.)

Interruptions in normal conversations are so common and apparently so important that we have tested their effects in a separate experiment. In brief, we found that if subjects do not have freedom to interrupt, they use fewer messages and more words per message. They also maintain a relatively constant number of messages in both speaking and typing.

Allowing subjects freedom to interrupt does not shorten the time needed to solve problems, nor does it result in any reduction in the number of words needed to reach solutions. What does happen is that when subjects have the freedom to interrupt, they package their words differently. They use more messages and fewer words per message, and they maintain a relatively constant message length whether they speak or communicate by typewriter. A final point of interest is that a subject is much more



MODE OF COMMUNICATION influenced the time required to solve problems. Here the average time taken by teams to solve problems is charted for 10 different modes of communication. The data fall into two fairly distinct groups. The faster five all involve the use of the voice in communication, whereas the voice is excluded in the five slower modes.

	COMMUNICATION-RICH	VOICE	HANDWRITING	TYPEWRITING EXPERIENCED TYPISTS	TYPEWRITING INEXPERIENCED TYPISTS
SOLUTION TIME IN MINUTES	29	33	53.3	66.2	69
NUMBER OF MESSAGES	230.4	163.8	15.9	27.2	31.5
NUMBER OF SENTENCES	372.6	275.9	24.9	45.8	44.1
TOTAL NUMBER OF WORDS	1,563.8	1,374.8	224.8	322.9	257.4
TOTAL NUMBER OF DIFFERENT WORDS	397.5	305.9	118.5	150.5	133.4
TYPE-TOKEN RATIO	.3	.3	.6	.5	.6
NUMBER OF WORDS PER MINUTE	190.3	171.2	17.3	18.1	10.2

EXPERIMENTAL RESULTS ***are enumerated for the solution of problems by various modes of communication. "Type-token ratio" is ratio of different words to total words. Problem solving by voice takes the least time but is wordier than the other modes are.***

likely to take control of a voice channel than of a channel or any combination of channels lacking the voice.

In considering the generality of our findings one might ask whether we have found certain general principles of human communication or have merely found out about the ways in which particular pieces of equipment are employed. We think we have found general principles. In our experiments we tested some of the various channels of communication in distinctly different ways. For example, in one experiment the voice channel was tested by having subjects converse through a cloth panel that was visually opaque but acoustically transparent. In another they conversed through a microphone and loudspeaker. In a third experiment each subject wore a microphone positioned a fraction of an inch in front of his lips. Similar variations were devised for handwriting tests. The most gratifying thing to me is that the results and the comparisons are almost identical under all variations of a given mode. In short, I believe we have been discovering general principles of communication (by voice, typing or handwriting) that are largely independent of the particular devices employed in mediating the communications. Another kind of evidence bearing on the generality of our findings is that the results obtained with our different modes of communication hold for all the different problems we have tested and for both job roles assigned to the communicators: source and seeker.

As so often happens in research, the questions our findings have raised are more numerous than the answers they have supplied. Every reader will have his own list of questions. Four in particular intrigue me.

First, how do communication patterns vary among different nationalities? Anthropologists and sociologists tell us (and our own experience seems to confirm) that communication patterns differ markedly among different cultural groups. Would we have obtained results of the same kind if we had studied people who speak a language other than English? What kind of concessions will telecommunication systems in general and computers in particular have to make for national and cultural differences?

Second, how do communication patterns vary according to the purpose of the communication? All our experiments so far have involved factual problems. The problems all had single answers, and the information needed to solve them was directed toward that one goal. Interactive communication, even with computers, may serve many different functions. A communicator may browse through masses of data for items that he needs or that merely excite his interest and curiosity. He may want to have briefings and status reports that will help him to anticipate weather conditions for the next few hours or days, review the state of the economy or update his information about the condition of patients in a hospital. Communications can also provide information necessary for reaching decisions among conflicting alternatives and can serve in argumentation, bargaining and persuasion. Virtually nothing is known about how communication patterns vary according to these diverse purposes.

Third, what happens to communication patterns as the number of communicators increases? Our experiments were all done with teams of two people, but communications often involve a group of people and perhaps a computer. The full implications of this kind of communication are not known.

Finally, what are the rules that govern normal human communication? Perhaps the most interesting question of all concerns the grammatical, syntactical and semantic rules that apply to such communication. In spite of the apparent unruliness of natural communication, it obviously follows some rules, because problems do get solved, often with surprising speed. How are meanings conveyed in natural conversations? How can we even go about investigating this problem?

This brief introduction to our program of research has conveyed only a few of the many interesting findings we have made. Still, it may be enough to excite the interest of others to try to understand what happens when people communicate. If enough people work on these problems, who knows? Perhaps at some future time you and I shall be able to find out about the latest developments in science not by reading articles such as the ones in this magazine but by conversing in ordinary English with a computer.