A double-blind test of astrology

from Shawn Carlson

Two double-blind tests were made of the thesis that astrological 'natal charts' can be used to describe accurately personality traits of test subjects.

ALTHOUGH there have been many published 'tests' of astrology, those with positive results (confirming the astrologers' thesis) have been largely dismissed by scientists on the grounds of technique. Those with negative results (disputing the astrologers' thesis) have been largely dismissed by astrologers on the grounds that they fail to test what the astrologers consider to be the essential aspects of their work. Indeed, astrologers complain that most scientific tests have tested the scientist's concept of astrology, not astrology as practised by the 'reputable' astrological community. Both criticisms may be valid.

My purpose has been to avoid these criticisms by designing an experiment that would meet the tight specifications of both the scientific and astrological communities. Such an experiment was designed with the help of scientists, statisticians and astrologers. We decided to test what we shall call (for simplicity) the 'fundamental thesis of natal astrology' as the proposition that:

> the positions of the 'planets' (all planets, the Sun and Moon, plus other objects defined by astrologers) at the moment of birth can be used to determine the subject's general personality traits and tendencies in temperament and behaviour, and to indicate the major issues which the subject is likely to encounter.

The device used by astrologers to make predictions is called a 'horoscope', in essence a picture showing the positions of the various astrological objects in the heavens on a backdrop of twelve equally-spaced imaginary sectors called 'houses', as seen from a particular place and time on Earth. Typically, a horoscope includes a table which shows the angular relationships (or 'aspects') between the astrological objects. If the place and time are those of a person's birth, the horoscope is called a 'natal chart' (see Fig. 1), from which astrologers derive information about a subject's personality and character. The descriptive text thus derived is called a 'natal chart interpretation'.

To satisfy both the scientific and astrological communities, we chose as advisers people held in high esteem by their respective communities. The astrologers helped us to formulate the proposition given above as central to 'natal astrology' (the subfield of astrology which deals with birth data) and yet scientifically testable. Care was taken to include all suggestions by the astrologers provided they could be followed without biasing the experiment

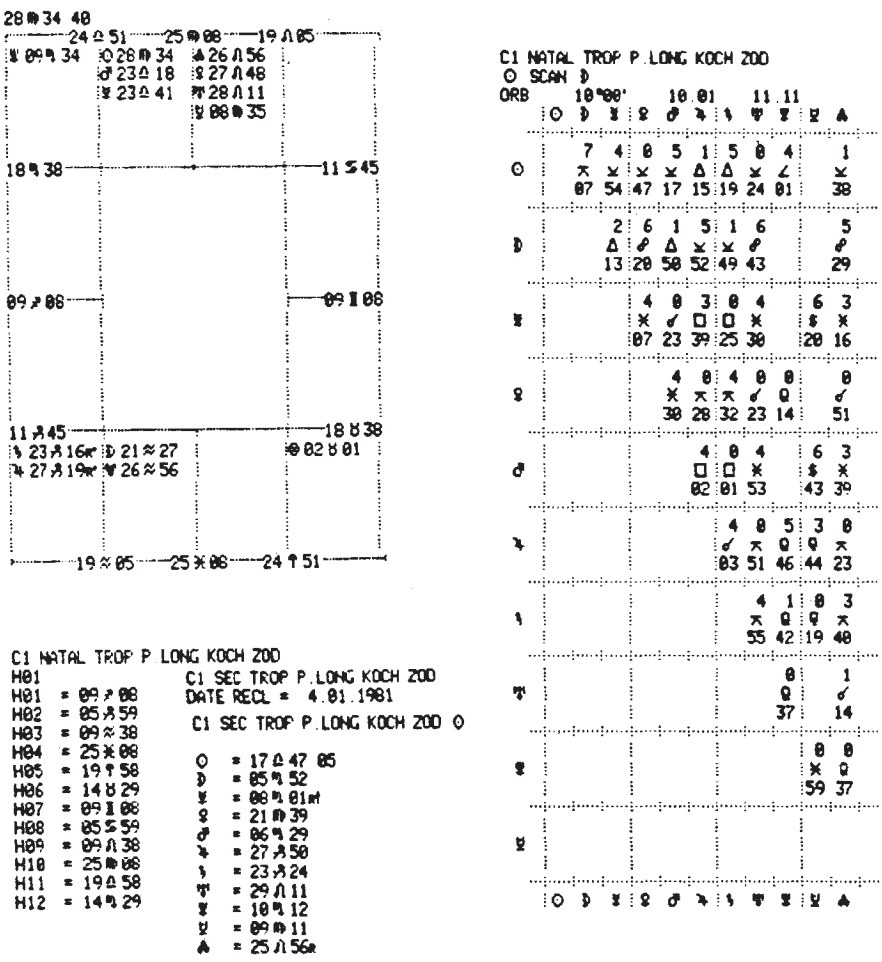

Fig. 1 Computer derived 'natal chart', showing positions of astrological objects as seen from the time and place of a person's birth.

for or against the astrological thesis. We took great care to eliminate all biases which could tend to 'randomize' the results and thus favour the scientific hypothesis over the astrological one. Similar care was taken to make sure that hidden clues were not available which could be used by astrologers or subjects to choose correct answers not based on astrological information alone.

The experiment designed by these means consists of two parts.

Part 1. Volunteers provided information from which their natal charts and interpretations were constructed by astrologers. Each subject then attempted to select his own natal chart interpretation from a group consisting of his own and two other interpretations chosen at random from the whole group. The subjects made first and second place choices; ties were not allowed. Subjects were also asked to rate each interpretation on a 1–10 scale (10 being highest) as to how well each interpretation fit them. If their selections are random (scientific hypothesis), we would expect them to select their own interpretation one third of the time. The astrologers predicted, given the design of the experiment, that the subjects would be able to choose their own interpretation "at least half" of the time.

Part 2. The participating astrologers were separately given the natal chart of a random subject and an objective and respected measure of his personality traits called the California Personality Inventory (CPI). They were also given two other

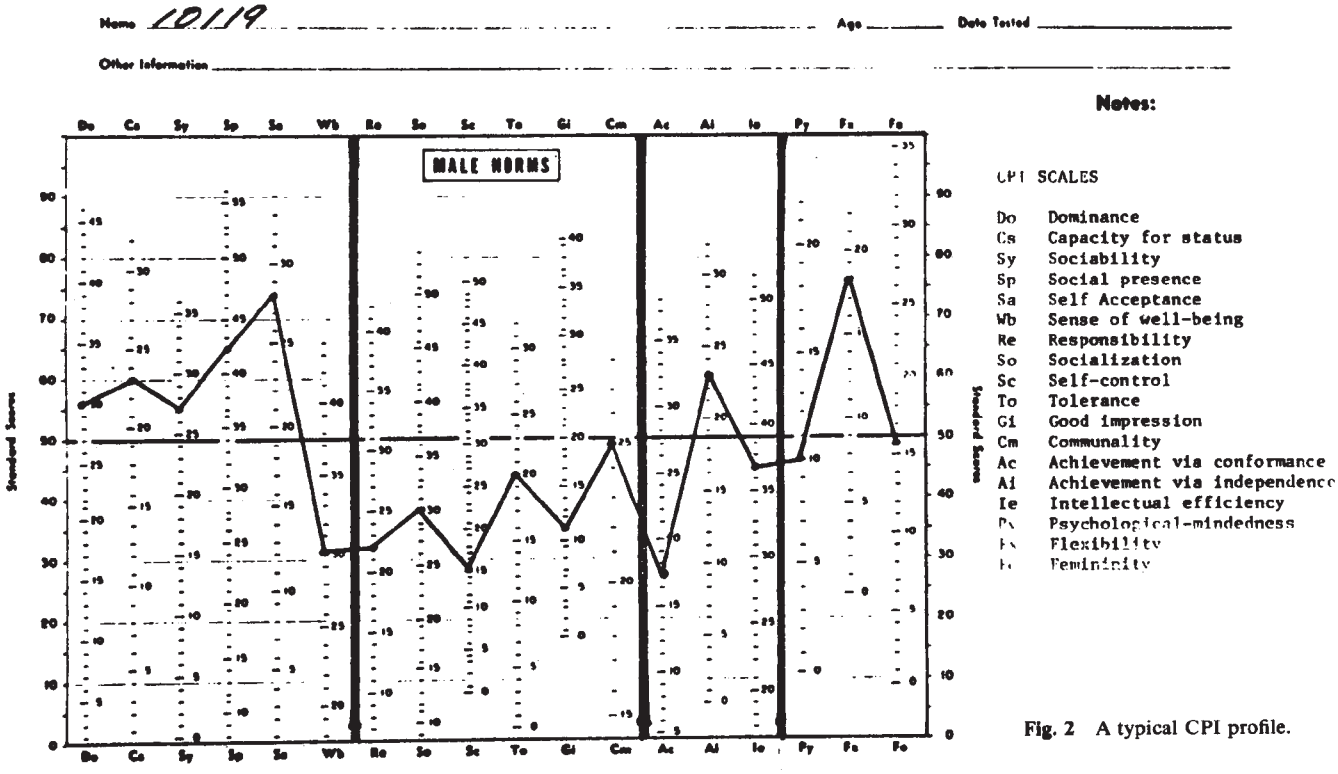

Fig. 2 A typical CPI profile.

CPIs chosen at random from the group of all the subjects' CPI test results. The astrologers were then asked to select the two CPIs (first and second choice, no ties allowed) which described personalities closest to the personality indicated by the natal chart. They also rated each CPI on a 1–10 scale (10 being highest) as to how closely its description of the subject's personality matched the personality description derived from the natal chart. The scientific hypothesis predicts a correct choice one third of the time; the astrologers predicted a correct choice half the time or more.

The two parts of the experiment are complementary; the first is not sensitive to biases in the CPI, the second does not make the possibly false assumption that subjects can accurately judge their own personalities. We decided at the outset to require a 2.5 standard-deviation increase over random chance to interpret the results as favouring the astrological hypothesis. Similarly, a disagreement by 2.5 standard-deviations or more would be required to reject the astrological hypothesis in favour of the scientific one. Otherwise, we decided in advance, we would drawn no conclusions of significance. No data were analysed until all the data had been collected.

Experiment design

To eliminate bias, both anticipated and unknown, we made extensive use of double blind techniques. All subjects were assigned a five-digit random code number. Neither the astrologers nor the experimenter knew what code number corresponded to which person. These lists were solely under the supervision of Richard A. Muller, Professor of Physics at the University of California, Berkeley.

Guidance was sought both from the scientific and astrological communities. To help ensure correctness of the testing method and statistical analysis, the scientific adviser was Professor Muller. So that participating astrologers should be respected by the astrological community, we sought the advice of the National Council for Geocosmic Research (NCGR), an organization which has been involved in much astrological research in the past and which has the respect of astrologers world wide; NCGR nominated persons who consented to be our astrological advisers, and who carefully reviewed the experimental design and made many suggestions.

After they were satisfied that the experiment was a 'fair test' of astrology, our astrological advisers established their predictions (50 per cent for both part one and part two) as the minimum effect they would expect to see. They also compiled a list of approximately 90 astrologers with some background in psychology who were familiar with the CPI and held in high esteem by their peers. It was the opinion of the advisory astrologers that a random sample from this list would be able to score at the predicted 50 per cent level. All were invited to participate; 28 accepted. (Only two astrologers who participated were not on the original list. They heard of the experiment and wanted to take part. After their qualifications had been vouched for by NCGR, they were admitted.)

Constructing a natal chart is a simple but laborious mathematical process, so computers are well suited for the task; several machines designed specifically for this purpose are available on the market. To save time and ensure accuracy, all natal charts were constructed by Mr. Caveney (President of the San Francisco chapter of NCGR) and Mr. Nelson (Secretary of the San Francisco chapter) on a Digicomp DR 70 Astrological Computer; they were spot-checked by hand calculation.

The California Personality Inventory (CPI) is a standard personality test[1–3] which has been used extensively since 1958. It was chosen over other available personality tests because the advising astrologers judged the CPI attributes to be closest to those discernable by astrology. By choosing this test we were thus trying to maximize the ability of the astrologers to match CPI data with natal charts without introducing a pro-astrology bias. Other experiments have been done using the CPI with apparently positive results[4].

The CPI consists of 480 true-false questions, each of which helps to rank a subject on one of 18 personality attribute scales (for example, dominance, passivity, femininity, masculinity). The subject's score on each scale is compared to the norm for that scale. The scores can be plotted on a graph (see Fig. 2) which readily conveys this information. Such a graph is called a 'CPI profile'.

Personality tests were graded, after names had been replaced by code numbers, by volunteers (undergraduate students) who were in no other way connected with the experiment. From spot

checks of the grading, we determined (95 per cent confidence level) that mistakes by the graders contribute an error of more than two points to CPI scores on fewer than 2.6 per cent of the individual scores, an insignificant effect.

Subjects were solicited by advertisements in San Francisco Bay area newspapers, classroom announcements and postings on and off the Berkeley campus. (To protect the confidentiality of the data and the rights of subjects, all procedures were checked by the University of California Office of Fair Treatment to Human Subjects before beginning data collection.) Approximately 70 percent of the subjects were college students and about one half of these were graduates. All subjects were required to fill out a questionnaire with their natal data (birthday, including exact time and location of birth). They were also asked whether they (1) believe in astrology, (2) believe somewhat, (3) have no opinion, (4) disbelieve or (5) strongly disbelieve in astrology, and to state whether they had ever had a natal chart constructed before. Subjects were not told that these questions affected subject selection, but those who chose "(5) strongly disbelieve" were eliminated, on the grounds that this opinion might bias them, either consciously or unconsciously, against selecting the interpretation which best fitted them. All those who had previously had a chart constructed were similarly eliminated because they might be able to select (or reject) the correct interpretation based on a knowledge of what to expect. Strong believers who had never had their charts done were, however, not eliminated; this belief alone could not help them select the correct interpretation. All subjects had to be at least 17 years old. Failure to take the CPI resulted in rejection.

To avoid the possibility that a subject may have had his natal chart constructed elsewhere, or may have changed his opinion about astrology, between the time he submitted his natal data and the time he was given the final natal interpretations (typically 8–10 weeks), all subjects were required to fill out a new questionaire before being asked to choose their own natal interpretations in Part 1 of the experiment. Two subjects were eliminated at this point, one who admitted to being a professional astrologer (and who had apparently lied on the first questionaire) and another whose opinion of astrology had changed from "disbelieve in astrology" to "strongly disbelieve in astrology".

We encouraged prospective subjects to participate by promising them a copy of their natal chart, CPI test results and interpretation, the completed natal interpretation, and a copy of the final results of the experiment.

Although we required a departure from random of only 2.5 standard deviations to interpret the results as favouring the astrological thesis, we originally planned to be able to distinguish between the two hypotheses at the four standard deviation level. This is why the number of subjects required was chosen to be 128, with a further 128 as a control group. Thus the total number of subjects was originally 256, but many of these did not complete all phases of the experiment.

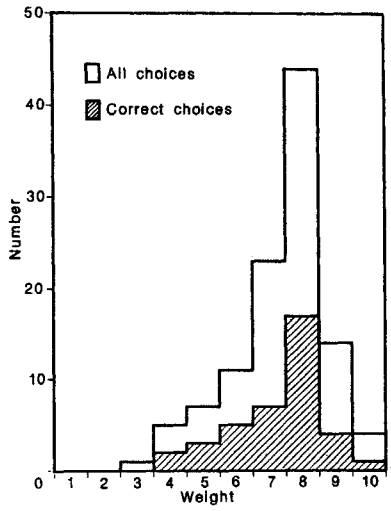

Fig. 3 Histogram showing weights assigned by astrologers to the CPI profiles they felt best fit the natal charts. CPI profiles rated higher are not more likely to be correct.

Many lost interest and did not return their data to us. Some moved in the time taken to send them the test materials and did not leave a forwarding address. Two room-mates became emphatically convinced that astrology was the work of the devil, and refused to continue in what they called 'an experimental test of evil'. We were forced also to eliminate 12 subjects either because they did not follow directions correctly or did not return all the needed materials to us.

The use of double blind techniques is most important during this stage of the experiment. During the process of rejection of data, the experimenter had no access to any information that might introduce bias. In the end, only 177 subjects (83 test group, 94 control) remained for Part 1 of the experiment. Neither were we able to collect all the data we had hoped to in Part 2 of the experiment. First, fewer astrologers than hoped for agreed to participate. 224 data envelopes were mailed to only 28 astrologers, some of whom simply refused to participate as promised. Some declined after they discovered how much time was required on their part. One tried to bargain his services in exchange for free access to our raw data, and declined to particpate when his terms were refused. For these reasons, we obtained only 116 usable subjects for Part 2 of the experiment. The large reduction in numbers was unanticipated and reduced the expected discrimination between hypotheses for Part 1 to 3.2 standard deviations, and for Part 2 to 3.9 standard deviations. However we do not believe that the loss of data could bias the results of the experiment in any significant way.

Bias and control

Experiments using human subjects are subject to a special class of biases which do not normally have to be considered by a physical scientist. An experiment must be designed so that the psychology of the subjects will not alter the results. The major potential biases which required specific control in the experiment design were as follows:

Sun-sign bias. By the astrological definition, the 'Sun-sign' refers to the constellation of the zodiac in which the Sun resides when the person is born. If the Sun-sign should play an important role in the average chart and if people are generally familiar with the characteristics of their Sun-sign (through newspaper horoscopes, for example), we might expect them to select the correct interpretation at a better-than-chance level regardless of whether or not the astrological hypothesis is correct.

To correct for this, each member of the test group was matched to a member of the control group born under the same Sun-sign. Following the astrologers' recommendations, we required that the age difference between these subjects be at least three years, so that their natal charts would be 'sufficiently dissimilar', but otherwise, the assignment was made randomly. Both test and control subjects were given the same three interpretations.

If the astrological hypothesis is false, members of both groups should identify the test subject's interpretation with equal frequency. If the hypothesis is true, the test group should score significantly higher than the control.

We believe that many other sources of bias are eliminated by the deisgn of the experiments and by the standard formats in which data was conveyed to and from the astrologers. The possibility that subjects (in Part 1) would be tempted to choose flattering interpretations (or the opposite) should not bias our conclusions, given that we are comparing hits and misses between test and control groups. We sought to eliminate from the interpretations clues to their origin that might guide a subject to a correct choice, such as astrological terms that might be part of subjects' general knowledge, and to eliminate from the data with which astrologers worked, information that might allow them, consciously or otherwise, to bury hidden clues in their interpretations. Thus we eliminated from the charts transmitted to the astrologers information about the time and place at which subjects had been born. (Although the charts in principle allow the reconstruction of this information, the process is time-consuming, and the outcomes unlikely to be of direct help to

astrologers.) We did not tell the astrologers of the gender of the subjects, partly because gender-specific elements of the interpretations might lead to bias, but also because the CPI scores cover different ranges for males and females.

To eliminate the possibility that subjects could pick up clues other than the astrological information we were testing, and to insure that the information given to subjects was as uniform, and thus as comparable as possible, the interpretations followed a predetermined format designed to specify what factors astrologers should derive from the chart and to set a limit on the length of written material. The format was developed in collaboration with the advising astrologers. The specific categories which astrologers were required to address were: (1) Personality/temperment; (2) relationships; (3) education; (4) career/goals; and (5) current situation. The astrologers typed each interpretation on pages supplied by the experimenter and containing the proper headings, again to keep the interpretations as uniform as possible.

The format also specifies: (1) that advice or predictions were not to be given, on the grounds that such information could not help the subject to select the correct interpretation but that it might well lead him to discard an accurate description because he disagrees with the advice or predictions given. (2) That no direct reference to the chart was to be made (e.g."you have sun in Leo"). (3) That no information relating to the subjects' ages was to be given.

Subjects were asked to rate each section of each natal interpretation on a 1-10 scale. Then they were asked to write down, for each section, the code number of the interpretation which fitted them best and second best.

One complication of our study arises because a subject's ability to select the correct description of himself from a given group must depend on how well he knows himself. If people generally have an inaccurate self-image, one would not expect subjects to select the correct interpretation no matter how accurate are the results of astrology. We devised the following scheme to understand this potential bias.

The CPI is generally accepted by psychologists as a moderately accurate description of a person's personality. Each test subject was given his own CPI profile and two others randomly selected from the group. He was then asked to select the profile which he felt best fitted him. Each subject was provided with the following: (1) three sample CPI profiles; (2) a synopsis of what high and low scores in each category tend to be for males and females; (3) a letter explaining about a CPI profile and how to go about making the selection.

To control for possible psychological bias, we elected to use the same test and control groups as in Part 1, but since the CPI is graded on different scales for males and females, we had to match male (female) test group members to male (female) control members. Thus, the test-control group assignments had to be re-established.

Scientists wishing to familiarize themselves with astrology as astrologers practise it may find the following works useful:

Gleadow, R. ***The Origin of the Zodiac*** **(Castle, Pasadena, 1968).**

Dean, G. ***Recent Advances in Natal Astrology*** **(Analogic, 1977).**

Rudhyar, D. ***The Astrology of Personality*** **(Doubleday, New York, 1970).**

Meyer, M. ***A Handbook for the Humanistic Astrologer*** **(Anchor, Manhatten Beach, California, 1974).**

Keyes, K. ***Master Guide to Preparing Your Natal Horoscope*** **(Parker, Los Angeles, 1984).**

Jones, M. ***The Sabian Symbols in Astrology*** **(Shambhala, Berkeley, 1953).**

George, L. ***A to Z Horoscope Maker and Delineator*** **(Llewellyn, St Paul, Minnesota, 1954).**

Gauquelin, M. ***The Cosmic Clocks*** **(Regnery, Chicago, 1967).**

Gauquelin, M. ***The Scientific Basis of Astrology*** **(Stein and Day, New York, 1966).**

Dean, M. ***The Astrology Game*** **(Beaufort, Beaufort, California, 1980).**

Obviously, since the natal chart depends entirely on the natal data, inaccuracies in the latter would produce inaccuracies in the former. The astrologers insisted that the birth time be accurate to within 15 minutes. In order to assure this, when the subjects took the CPI, they were obliged to show documentation of their natal data including, especially, birth time. Although we preferred birth certificates, hospital and county record or other 'official' documentation, we also accepted baby books provided that the birth time was recorded when the child was born. (A more complete discussion of biases may be found in L.B.L. Preprint No. 20480.)

Double-blind procedures

An important difference between this and many previous tests is our extensive use of double-blind techniques. The important procedures we followed are outlined below.

As soon as subjects had taken the CPI, their questionnaires were put into alphabetical order and given to an assistant, who assigned a five random-digit code number to each in turn. No two subjects were assigned the same code number. The assistant then filled out three 3×5 index cards for each subject. On the first he put Name-Code Number; these he filed in alphabetical order. On the second he put Code Number-Name; these were filed in numerical order. The purpose for these cards was for easy record-keeping. The cards were maintained under the supervision of Prof. Muller, and could be released to the experimenter only with his consent. At no time during the data collection did the experimenter have access to any information relating subjects' identities to code numbers. This control was abandoned only when all the data had been collected and the methods of analysis had been established.

The assistant also made a third set of cards each containing the code number of a particular subject and his natal data. These cards were given to astrologers Michael Caveney and Chris Nelson in envelopes unopened by the experimenter, for the construction of natal charts. Since the CPI was given at three different sessions (typically about 4 weeks apart), not all the assignments were made at the same time, nor did the astrologers receive all the natal data cards at one time. They did receive, however, the natal data cards within five days of the CPI testing dates.

Subjects who failed to show documentation of birth time, date and location when they took the CPI were automatically put into the control group (there were 43 such subjects). All questionaires in each group were sorted into twelve Sun-sign groups. We then randomly assigned subjects to the control group until the number of control group members equalled the number of subjects remaining in each Sun-sign group. The remaining subjects comprised the test group. If there was an odd number of subjects in a Sun-sign group, the odd person was put in the control group.

The questionnaires in each Sun-sign group were then shuffled thoroughly and the names and birthdates were listed in the order in which they landed. As stated earlier, we required the birth dates to be at least three years apart to ensure that the natal charts were sufficiently dissimilar. The first test subject of each Sun-sign group was then matched to the first control subject of the same group whose birth date was far enough apart. The procedure was repeated for the second, third, etc. test subjects until we were left with test and control subjects whose birthdates were within three years of each other (which happened in all 12 Sun-sign groups).

To include these remaining subjects, we were forced to do some rematching in the

Table 1 Data from subject selections of natal chart interpretations

	Total	First choice	Second choice	Third choice		
Test group	56	25	16	15	18.67±3.53	CPI PROFILE
Control group	50	21	13	16	16.67±3.33	SELECTION
Test group	83	28	33	22	27.67±4.29	INTERPRETATION
Control group	94	42	34	18	31.33±4.57	SELECTION

following way. We started at the top of the control group and went down until we found a birthdate which satisfied the requirements that it was at least three years away from the unmatched test person and that the birthdate test group member to which it was attached was at least 3 years away from one of the unmatched controls. We then rematched the originally-matched control group member with the previously unmatched test subject, and matched the unmatched control group member with the previously matched test member. This was continued until all the test and control group subjects were matched. Because seven Sun-sign groups had an odd number of subjects, there remained seven unmatched control group members.

For the second part of the experiment, male (female) test members had to be assigned to male (female) control members. Thus, the above test subject to control subject matching had to be redone. Since the test and control groups were randomly chosen, no changes in them were made. However, all the matchings changed as male (female) test subjects were randomly matched to male (female) control subjects.

During the conduct of the experiment, each subject was given two envelopes, one containing the materials needed for the selection of the natal chart interpretation and the other containing the materials for the selection of the CPI profile. The natal interpretation envelope contained: (1) three natal interpretations; (2) a pretyped sheet on which the subjects were to detail their choices; (3) a questionnaire asking their opinion of astrology and whether they had had their chart done before (See Experiment Design); and (4) a letter explaining how they were to go about making the selections. The CPI envelope contained: (1) three CPI profiles; (2) a preformatted sheet on which the subjects were to detail their first and second place choices, plus 1-10 ratings; (3) a summary of what the CPI scores mean; and (4) a letter explaining how they were to go about making their selections. Since we had labelled the natal charts with the code number of the person for whom they were constructed, we could not do the same with the CPIs, since the subjects might recognize a code number appearing twice. The CPIs were first labelled with the code numbers of the persons to which they corresponded, and then relabelled by finding $\ln(1/x)$ for each code number, x, on a hand calculator and taking the last 5 digits on the calculator display as the new code number. All natal interpretations and CPI profiles were put in numerical order in the envelopes.

The astrologers were sent materials in two separate mailings. In the first, each received: (1) the number of natal charts he had agreed to interpret when he chose to take part (typically 4); (2) a copy of the format by which the charts were to be interpreted; (3) the paper with headings on which the interpretations were to be typed; (4) a letter explaining the symbols used on the computer-constructed natal charts, deadlines and notice of when they might expect the materials for part two of the experiment; (5) a postage-paid return envelope. After they had returned the natal charts and interpretations, the astrologers received the second mailing containing: (1) the number of natal charts (plus three CPIs for each natal chart) which they had agreed to match to CPI profiles; (2) a copy of *The Interpreter's Syllabus for the CPI*[2] (a booklet explaining all the CPI attributes and how to interpret them in detail); (3) a preformatted sheet on which the astrologers were to detail their first and second choices and ratings; (4) a letter explaining how to go about making the choices; (5) a postage-paid return envelope for the data.

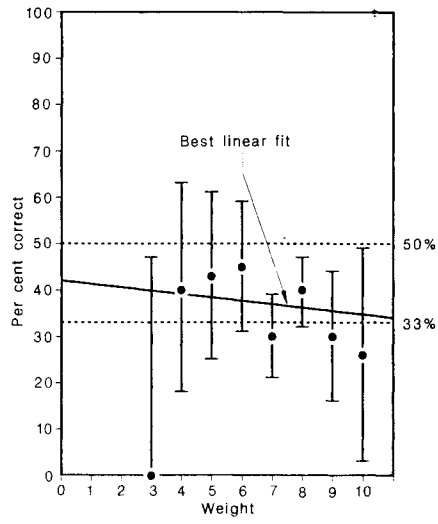

Fig. 4 Graph showing percentage correct versus rating for astrologers first place choices in CPI profile natal chart matching. The best linear fit is consistent with the scientifically predicted line of zero slope. No significant tendency for the astrologers to be more correct when they rate a CPI as highly matching a natal chart.

To save the astrologers work, they were allowed to make the CPI matchings to the natal charts they had already interpreted. They were also typically sent an additional natal chart and CPIs to match. To ensure that the astrologers would not mix up the CPI profiles between natal charts, the three profiles for each natal chart were ordered randomly, then labelled *a*, *b* and *c* respectively.

A total of 226 natal charts were sent out to be matched with CPI profiles. Of these, only 15 had no documentation of birth time. None of these 15 were returned by the astrologers to be included in our data.

Results

Part 1. *Subject selection of natal chart interpretations.* After a predetermined data collection period of 10 weeks, we had astrological data from 83 test and 94 control subjects and CPI self-selection data from 56 test and 50 control subjects. The data are displayed in Table 1.

The test group selected the correct interpretation as its first choice at the rate of 0.337 ± 0.052, the control group at the 0.447 ± 0.049 rate, 2.34 standard deviations above chance. Although this fluctuation is less than 2.5 standard deviations, the level we had chosen to call 'significant', it does require comment. Since this fluctuation occurred in the control group and control subjects were not given their own interpretations, this cannot be interpreted as a possible astrological effect. Neither can it be correctly attributed to Sun/sign bias, since the test group did not score near the same level. We thus interpret this as a statistical fluctuation. The test group chose the correct interpretation as second best, describing them at the 0.398 ± 0.052 rate while the control group did so at the 0.362 ± 0.049 rate. Finally, the correct interpretation fell as the test group subjects third choice at the 0.265 ± 0.052 level and, for the control group, at the $0.191 \pm$

Table 2 Data from astrologers matching natal charts to CPI profiles

	Total (n)	Chance ($n/3$) [Expected s.d.]	Astrologers predicted ($n/2$) [Expected s.d.]	No. of correct CPI chosen	Standard deviation away from 0.35	Standard deviation away from 0.50
First choice	116	38.5 [5.1]	58.5 5.4	40	+.256	3.34
Second choice	114	38.0 [5.0]	None	46	+1.48	—
Third choice	114	38.0 [5.0]	None	28	2.0	—

The data are consistent with chance, inconsistent with astrological hypothesis.

0.049 rate. All this is consistent with the scientific hypothesis.

When the first few data envelopes were opened, we noticed that on any interpretation selected as a subject's first choice, nearly all the subsections were also rated as first choice. We then realized that we had no way of guaranteeing that subjects were rating each section of the interpretations independently of others they had already read. Without such a guarantee, spurious results favouring either hypothesis could have easily appeared. So we rejected these data as not having been collected under the proper controls.

Next we looked to see how well the subjects were able to select the correct CPIs. The test group selected the correct CPI as their first choice at the 0.446 ± 0.063 rate, while the control did so at the 0.420 ± 0.066 rate, showing no significant difference between the two groups. The test group chose the correct CPI as the second place choice at the 0.286 ± 0.063 rate, while the control did so at the 0.260 ± 0.066 rate, again no significant difference between the two groups. Finally, the test group chose the right interpretation as their third choice at the 0.268 ± 0.063 rate while the control group did so at the 0.320 ± 0.066 rate. Once more, there is no significant difference between the two groups.

Part 2. *Natal Chart-CPI matching.* A total of 116 data envelopes were returned by the astrologers. In the 116 envelopes there was a total of 116 first place choices, 114 second place choices and 320 ranked choices (weight factors indicating how well the astrologers felt the natal chart matched each CPI, on a scale of 1 to 10) (Table 2).

The data were first analysed without taking the 1-10 weight factors into account. The astrologers selected the correct natal chart as their first place choice at the 0.34 ± 0.044 rate, in agreement with the scientific hypothesis of 0.33 and in disagreement with the astrological hypothesis of 0.5 by 3.3 standard deviations. The correct CPI was chosen as the second place choice at the 0.40 ± 0.044 rate which is also consistent with the scientific hypothesis. (The astrologers had made no firm prediction about the second place choice.) The correct CPI was chosen as the third place choice at the 0.25 ± 0.044 rate, again consistent with the scientific hypothesis.

Next we took the weights into account, by a method established before studying the data. (The establishment of methods before data analysis is important in order to prevent the subtle bias that comes from selection of analysis procedures.) We first made a histogram of the weights which the astrologers assigned to all their place choices, regardless of whether or not those choices were correct (Fig. 3). The data are sharply peaked at a weight of about 8. The second histogram in Fig. 3 shows the ratings of only those first place choices which

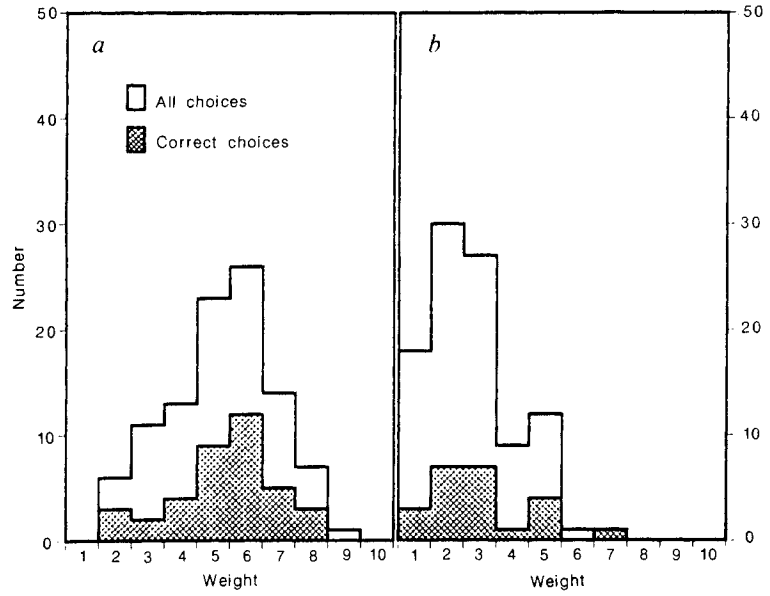

Fig. 5 Histograms showing rates assigned to astrologers' second (*a*) and third (*b*) place choices.

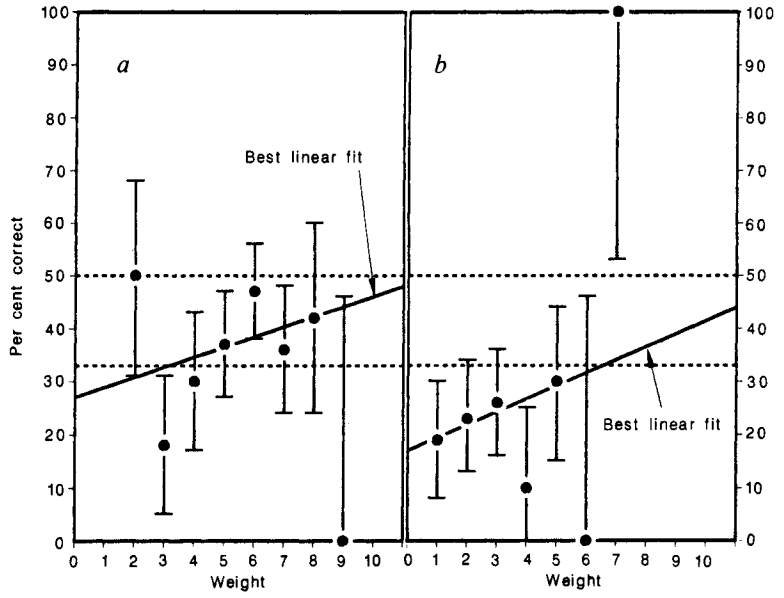

Fig. 6 Percentage of correct CPI profiles versus rates, chosen by astrologers as their second (*a*) and third (*b*) place choice. Best linear fits are consistent with chance.

were correct. If the astrological hypothesis were true, one might expect the correct first place choices to have higher weights on average than the whole group of first place choices. Thus, the new histogram should be skewed to the right. On comparing the two histograms, however, we see that they are very similar; no such skewing appears.

The scientific hypothesis predicts that 1/3 of the choices at any weight should be correct choices. Figure 4 shows the percentage correct for each weight with the appropriate error bars, and the best linear fit with slope -0.01 ± 0.02. The slope is consistent with the scientific prediction of zero slope. The same analysis on the second and third place choices yields Figs 5*a*, *b* and 6*a*, *b*. The slope of the best linear fit to the data on Fig. 6*a* is 0.019 ± 0.02 while that of Fig. 6*b* is 0.0026 ± 0.02 both consistent with the scientific hypothesis (zero slope).

Conclusions

From the results of Part 1 (subjects selecting interpretations), we notice that the test group scored at a level consistent with chance and within 2.5 standard deviations of the control group. The large (2.34 s.d.) but not significant (less than 2.5 s.d.) fluctuation in the control group is attributable to statistical fluctuation, not to a Sun-sign bias. These results are consistent with the scientific hypothesis. However we cannot use the result to rule against the astrological hypothesis, because the test subjects were also unable to select their

own CPI profile at a better-than-chance level. At the 95 per cent confidence level, the test subjects were unable to select their own CPI profile at better than the 0.57 rate. There are many reasons which could explain why the test subjects were unable to select the correct CPIs at a higher rate:

(1) Subjects may have had difficulty relating to the graphical presentation of the CPI information.
(2) Some subjects may have recognized correct information about themselves, but subconsciously chose a CPI which did not describe them as well to avoid admitting they have certain character traits. Such denial in a large percentage of the subjects would tend to cancel a positive effect.
(3) The CPI may not test the kind of attributes by which subjects may easily recognize themselves.
(4) People may be unable to recognize accurate descriptions of themsleves.

Our experiment does not distinguish between these possibilities. Professor H. Gough (author of the test and respected experimental psychologist) is familiar with nearly all published experiments using the CPI. At our request he searched through the literature for any experiment demonstrating the ability of test subjects to recognize accurate descriptions of themselves. To his and our knowledge, no other test of this kind has ever been done. Thus, we believe there exists presently no scientific evidence from which one can conclude that subjects can select accurate descriptions of themsleves at a significant rate.

If subjects cannot recognize accurate descriptions of themselves at a significant level then the experiment would show a null result no matter how well astrology worked. On the other hand, any astrological effect demonstrated in this way would require a consistency check. One would have to see if subjects could recognize the kind of information astrologers give them about themselves, which was derived in a manner *known* to be reliable. Thus, until and unless such a self-recognition ability can be shown, we conclude that subject selection of astrologically derived information is a poor test of astrology. (This is a problem in approximately 30 per cent of all experiments which claim a significant astrological effect.)

The conclusions to be reached from Part 2 (CPI-natal chart matching) of the experiment are somewhat more illuminating. What is striking about these data is how poorly the astrologers performed, when their performance is compared to their predicted rate. It is consistent with chance, and is at the very significant 3.3 s.d. level below the astrologers' prediction. This is well beyond the 2.5 s.d. requirement we established before the beginning of the experiment as sufficient to refute the astrological hypothesis.

Before the data had been analysed, we had decided to test to see if the astrologers could select the correct CPI profile as either their first *or* second choice at a higher than expected rate. The scientific hypothesis predicts the CPI will fall in the first or second choice 66 per cent of the time. The astrologers did not make a specific prediction as to what they expected the rate to be. If the correct CPIs are chosen in the first and second place choices, then they will be depleted from the third place choice. Since the rate at which the astrologers chose the correct CPI as their third place choice was consistent with chance, we conclude that the astrologers were unable to chose the correct CPI as their first or second choices at a significant level.

In Fig. 4 the data are clearly inconsistent with the 'at least' 0.5 level predicted by the astrologers. Nor do the data suggest that the astrologers are more likely to be correct when they rate a CPI as well fitting the particular natal chart than they are when they weight it as poorly fitting the natal chart. The data appear randomly scattered about the 0.33 line and is hence consistent with chance. The scientific hypothesis predicts a line of zero slope, consistent with the slope observed. Figures 5 and 6 likewise show no convincing evidence that the astrologers tended to rate the correct CPIs higher than the incorrect CPIs.

We are now in a position to argue a surprisingly strong case against natal astrology as practised by reputable astrologers. Great pains were taken to insure that the experiment was unbiased and to make sure that astrology was given every reasonable chance to succeed. It failed. Despite the fact that we worked with some of the best astrologers in the country, recommended by the advising astrologers for their expertise in astrology and in their ability to use the CPI, despite the fact that every reasonable suggestion made by the advising astrologers was worked into the experiment, despite the fact that the astrologers approved the design and predicted 50 percent as the 'minimum' effect they would expect to see, astrology failed to perform at a level better than chance. Tested using double-blind methods, the astrologers' predictions proved to be wrong. Their predicted connection between the positions of the planets and other astronomical objects at the time of birth and the personalities of test subjects did not exist. The experiment clearly refutes the astrological hypothesis.

I thank Professor Richard Muller for providing funding from his Alan T. Waterman Award, and for invaluable assistance and guidance; Professor Doksum, Statistics, UC Berkeley, for acting as an advisor at the conception of the experiment; Tony Joseph for many helpful conversations and for being our astrological advisor; Michael Caveney and Chris Nelson for constructing the natal charts and being astrological advisors; the participating astrologers for their expertise, time, and effort; Terry Mast, Teresa Weed, Geoffrey Dean for helpful criticisms; Kent McArthur who first suggested the use of personality tests; Susan Miller, who assisted in the advertising, and to those Berkeley undergraduates who gave up their time to act as assistants and CPI test graders.

Shawn Carlson is at the Department of Physics, University of California, Berkeley, California 94720, USA. Present address: Department of Physics, University of California, Los Angeles. Address for correspondence: Bldg 50, Rm 348, Lawrence Berkeley Labs, Berkeley, California 94720, USA.

Received 11 April 1983; accepted 14 October 1985.

1. Gough, H. G. *Manual for CPI* (Consulting Psychologists Press: Palo Alto, 1957).
2. Gough, H. G. *Advances in Psychological Assessment* Vol. 1, (ed. McReynolds, P.) (Science and Behavior Books, Palo Alto, 1968).
3. Megargie, E. I. *The CPI Handbook* (Jossey-Bass, San Francisco, 1972).
4. Pellegrinli, R. J. *J. Psychol.* **85**, 21–28 (1973).