

A Scientific Inquiry Into the Validity of Astrology

JOHN H. MCGREW and RICHARD M. MCFALL

Psychology Department, Indiana University, Bloomington, IN 47401

Abstract—Six expert astrologers independently attempted to match 23 astrological birth charts to the corresponding case files of 4 male and 19 female volunteers. Case files contained information on the volunteers' life histories, full-face and profile photographs, and test profiles from the Strong-Campbell Vocational Interest Blank and the Cattell 16-P.F. Personality Inventory. Astrologers did no better than chance or than a nonastrologer control subject at matching the birth charts to the personal data; this result was independent of astrologers' confidence ratings for their predicted matches. Astrologers also failed to agree with one another's predictions.

Nearly one-third of the population in Western countries believes in astrology; another third cares enough to attend to astrological predictions at least some of the time (Eysenck, 1982). Most scientists, in contrast, have simply ignored, dismissed, or denounced astrology (Bok, Jerome, & Kurtz, 1975). As a result, the scientific community has been vulnerable to accusations of dogmatism and authoritarianism (Eysenck, 1982; Rockwell, Rockwell, & Rockwell, 1978).

The central claim of astrology is that the psychological attributes and personal destinies of individual human beings are related to the positions of heavenly bodies at the moment of each person's birth. The fairest test of astrology assesses the accuracy of predictions made by qualified astrologers on the basis of global interpretations of complete horoscopes (Eysenck, 1982) (e.g., not simply limited to sun signs). Unfortunately, many of the studies that have attempted such a test (Clark, 1961, 1970; Dobyns, 1976; Gauquelin, 1973, 1978; Joseph, 1975; Vidmar, 1979) have suffered from methodological problems (e.g., small sample sizes) that have prevented them from providing unambiguous evidence concerning astrology's validity (Eysenck, 1982). For example, subjects providing the personal and birth information sometimes have not been blind to the fact that they were participating in astrology research (Vidmar, 1979); thus, the possibility of subtle biases both in subjects' selection and in subjects' self-reports cannot be ruled out (Mayo, White, & Eysenck, 1978). Also, the studies have tended to focus on astrologers' ability to use birth information to predict only single dimensions of personal information (e.g., occupation). A more ecologically valid test would be to assess how well astrologers could match specific birth information to the full, complex pattern of each individual's life experiences and personal characteristics. Finally, because almost none of these studies was

designed in close association with astrologers, the studies have been vulnerable to the criticism that the test was a flawed or unfair representation of standard astrological practice.

To counter these potential criticisms, it has been suggested that the best procedure for testing claims about astrology would be a cooperative one, in which astrologers and scientists jointly designed a test of the hypotheses. To our knowledge, only one such study has been reported previously (Carlson, 1985). Unfortunately, the results of that study were not unambiguous, both from the point of view of methodology and astrology. The study consisted of two parts. In part one, nonastrologer-test subjects exhibited equal difficulty in picking from a group of three possible interpretations (the interpretations were prepared by astrologer subjects) the horoscope interpretation that best described themselves, and in picking from a group of three possible profiles the California Personality Inventory (CPI) profile that best described themselves. The part one data permitted no clear conclusions about the relative accuracy of personality descriptions derived from astrology. In part two, the astrologer subjects were unable to pick from a group of three possible profiles the correct CPI profile of test subjects (for whom they had the corresponding horoscopes). The author concluded that the data from part 2 made, "a surprisingly strong case against natal astrology. . . ." (Carlson, 1985, p. 425).

This conclusion may be premature. Although it is true that the astrologers could not select the CPI profile corresponding to the test subjects, it also is the case that the test subjects themselves could not select their own CPI profiles. That is, both the astrologers and the test subjects failed on the identical task! Because the test subjects' failure on the task could not have been due to the invalidity of astrology, their failure must have been due to some non-astrological difficulty with the task; thus, the astrologers' failure on the task may have resulted from the same nonastrological difficulty (e.g., difficulty understanding the personality terms or profile configurations used to describe individuals' traits on the CPI). The experimental procedure cannot distinguish between these rival hypotheses. Given this methodological inadequacy, we cannot conclude that the astrologers' failure was due to the invalidity of astrology. Thus, the results of this study are inconclusive.

Two further criticisms of this study, and other studies of its type, are pertinent from an astrological perspective. First, the information derived from standard psychological tests, like the CPI, may not include the types of information that astrologers require to complete a matching task successfully. In designing the present study, the astrologers from the Indiana Federation of Astrologers (IFA) were asked to generate a list of the kinds of personal information that they would require to perform the matching task accurately. No limits were placed on the kind of information this might include. The astrologers generated a set of 61 questions that covered an extremely broad range of information. Based on this list, the first author, in collaboration with the IFA representatives, developed a 61-item questionnaire, the Personal Characteristics and Life History Summary (PCLHS). The PCLHS

asked questions about hobbies, interests, religious beliefs, physical characteristics, personal talents and achievements, family background, dates of parent or sibling deaths, dates of moves across the country, health problems, attitudes toward authority, sex and commitment, pet peeves, favorite colors, punctuality, dependability, and variations in the personal energy cycle. Neither the CPI nor any other standard psychological instrument contains this type of information, yet the astrologers considered this information vital to be able to perform the experimental matching task accurately.

Second, previous studies of astrology typically have been limited to assessing astrologers' ability to predict personality traits, even though practicing astrologers as often use astrology to predict the occurrence and timing of specific events in the life of an individual (e.g., if and when events may occur, when clients will perform at their best). In the present study, the astrologers specifically requested information about the occurrence and timing of important events in the lives of the test subjects; for example, they wanted information on subjects' "personal energy cycles," dates of moves across the country. They felt that both categories of information—personality traits and life events—would be critical to performing the matching task successfully. Although the present study did not analyze these two information categories separately, it did provide both to the astrologers, thus, providing an indirect test of astrologers' ability to make predictions about the occurrence and timing of events.

The present study, then, was designed to overcome the methodological and "astrological" limitations of previous studies. It was conducted with the full and close cooperation of the Indiana Federation of Astrologers. The final experimental protocol was adopted only after it had been approved by the IFA as a fair and reasonable test of the predictive capabilities of astrologers.

Before the IFA agreed to collaborate on this study, there was a protracted negotiation period. The astrologers, understandably, were wary of becoming involved with research that might be biased against them or that would provide no opportunity for success. Initially, the first author gained entry through the influence of a well-known numerologist with whom he was friendly and who vouched for his integrity. Several letters were exchanged, in which the first author expressed his genuine desire to investigate astrology fairly and without prejudice. Finally, a horoscope of the first author was prepared by the IFA, both to inform the first author of the kinds of information available from a horoscope, and to provide the IFA with astrological evidence of his sincerity. After this final step, the IFA agreed to sanction the project.

The design of the project proceeded in iterative steps. As was mentioned previously, the astrologers were encouraged to determine the information necessary to complete the matching task. During this process, the astrologers placed several restrictions on the volunteers and asked for nonstandard data sources. For example, the astrologers asked for photographs of the volunteers in order to determine astrological body types. Also, the astrologers

asked that the subjects be restricted to age 30 or older, because they believed that younger subjects would not have manifested the mature personality characteristics reflected in the horoscope. Similarly, the astrologers asked for, and we were able to provide, a diverse sample of people, ranging in occupational status from professional to blue collar. The diversity included: a former prostitute, a lawyer, a never-do-well politician's son, a "bum," an entertainer, a journalist, a sailor, and a fireman. The astrologers even were involved in the choice of formats for the questionnaire; they felt that open-ended questions would better represent astrological practice.

As this illustrates, the design of the project was completely collaborative. The astrologers were cooperative, appropriately directive, and enthusiastic about the project. They gave freely of their professional time and conducted their part of the experiment in good faith. The resulting experiment was as close as we were able to come to a consensually fair, methodologically sound, and astrologically achievable design.

Method

Subjects

The experimental subjects were six individuals (one man, five women) nominated by the IFA as astrologers with superior ability. Superior ability among the subject group was documented in several areas. One of the astrologers authored and published two books on various aspects of astrology and published a national newsletter on astrology. Another had been a professional astrology writer for a syndicated column. All of the astrologers had at one time or another been professional counseling astrologers (earning money by the practice of astrology).

Although the principle experimental question was whether the astrologers could do better than chance in their predictions, one control subject, a male graduate student in clinical psychology, was included in the study. Finally, 10 college students (7 men, 3 women) enrolled in the introductory psychology course at Indiana University were recruited as subjects in a separate experimental control task.

Design

Each of the six astrologers and one control subject was given two sets of materials, both pertaining to the same 23 individual test cases. One set contained personal information about each test case; the other set contained detailed birth information on each case. Each astrologer subject also received natal charts (i.e., horoscopes) for each volunteer; these had been prepared previously by the IFA solely on the basis of each volunteer's birth information. The order of cases within each set had been randomized. The subjects' task was to match the personal information to the corresponding birth information for each of the 23 test cases.

Case Materials

As was mentioned previously, the IFA had generated a list of the types of personal information that the astrologers would need from each person serving as a test case. This list was used to derive the PCLHS, which was included in the case materials.

In addition to the PCLHS, two standardized psychological tests were selected, the Strong–Campbell Interest Inventory—Form T325 and the **Cattell 16 P.F.—Form B**, to provide information concerning general interests, potential vocations, and personality traits.

At the request of the IFA, two photographs—frontal and profile—were taken of each test-case person, to be used to determine "astrological body types."

Finally, the exact date, time, and place of birth was obtained for each test-case person. Subjects were asked to obtain their exact birth time—accurate to within 10 minutes—and to verify the birth time with the birth certificate, hospital records, or county records.

Collection of Test Case Materials

Twenty-three Caucasians (4 men, 19 women) were recruited to provide the test-case materials for this study. These volunteers responded to an announcement in the local newspaper offering free vocational testing to native-born American adults, 30 or 31 years of age. The age range was restricted so that the personal information and birthdates could not be matched simply on the basis of age-related cues in the photographs.

At the time of scheduling for the individual testing sessions, volunteers were asked to obtain and bring with them accurate information concerning the date, place, and time of their birth. The cover story for this request was that the experiment concerned the possible influences of maternal diurnal cycle on the physical condition of the mother during the birth process and on the later development of the neonate. For example, early morning births may alter the mother's hormone levels by disrupting her normal sleep cycle.

In the testing session, each volunteer signed a consent form, completed the personal-information measures and questionnaires, and supplied the requested information concerning the precise date, time, and place of birth. Only one volunteer had not brought the requested birth information; she provided it later by telephone after consulting her birth certificate. Finally, frontal and profile Polaroid photographs were taken from a standard distance of one meter. The volunteers were not informed of the true nature of the study until after they had completed all of the measures; however, two volunteers reported during debriefing that they had suspected that the study might be related to astrology.

Each volunteer's PCLHS was edited to remove any information that might be linked directly or indirectly to the time or place of the person's birth. Only four cases required minor editing. Then, for each volunteer, a

personal information file was constructed consisting of the PCLHS, a computer-scored summary of the Strong-Campbell Interest Inventory, a personality profile derived from the 16-P.F., and Xeroxed copies of two photographs. Each case file was identified by the last four digits of the volunteer's social security number. Recorded on a separate coded list, with no other identifying information, were the precise date, time, and place of each volunteer's birth.

Procedure

Each subject (six astrologers; one control) received the case materials, a separate list of the volunteers' birth dates, times, and places, and an answer sheet. The astrologers were instructed to work independently. Subjects were to record their choices on the answer form. Next to each birth date, they were to enter the code number of the case file that was their first choice for a match. Next to that choice, they were to indicate their confidence level on a scale ranging from 0 to 100 percent. If they wished, they could record alternative choices beside their first choice. Subjects were not limited on the number of alternative choices they could provide for any one birth date. Also, they could provide alternatives for every birth date or only for some of the birth dates. Confidence ratings were not obtained for alternative choices. The astrologers mailed their completed answer sheets directly to the experimenter.

To rule out the possibility that the test cases and birthdates might be matched simply on the basis of age-related cues contained in the photographs, a separate control study was conducted. Ten undergraduates recruited from the introductory psychology subject pool were scheduled for individual sessions in which they were asked to rank order the 23 test-case volunteers from youngest to oldest solely on the basis of their Polaroid frontal photographs. These photo rankings were compared to the actual age ranking of the test cases. Across the ten judges performing this task, the number of exact matches between photo ranks and age ranks ranged from zero to three, with a median of zero. To have done better than chance, with alpha set at 0.05, a judge would have needed at least four correct matches out of the 23 test cases (Feller, 1961). Thus, these results indicated that the test cases could not be matched reliably to their birthdates solely on the basis of age-related cues in the photos.

Results and Discussion

Examination of subjects' first-choice attempts at matching the test cases to the birth information revealed that the number of correct matches by the six astrologers ranged from zero to three, with a median of one. The control subject achieved three correct matches, thus equalling the most successful of the astrologers. In short, no subject—astrologer or control—performed the matching task at a level that was significantly beyond chance (Feller, 1961).

There was little relationship between astrologers' confidence in their predicted matches and the accuracy of their predictions. Overall, the astrologers seemed confident in their predictions, with a mean confidence rating of 73.5%. Across astrologers, the Pearson correlation between number of accurate predictions and mean self-rated confidence was nonsignificant ($r = .03$). For those astrologers who had at least one correct match, the confidence ratings for correct and incorrect predictions (means = 76.4 and 72.8, respectively) were not significantly different ($t = .473$, $df = 90$).

The number of second choices offered by the six astrologers were: 0, 0, 2, 6, 8, and 21, respectively. The control subject offered 2 second choices. When subjects' second-choice predictions were substituted for their incorrect first choices, the number of correct matches increased for only two subjects. The astrologer who had offered 21 second choices increased from two to three correct matches. The control subject increased from three to four correct matches. Thus, when second choices were considered, it was still the case that no astrologer performed better than chance, but now the control subject achieved more matches than any astrologer.

At the very least, if astrology constitutes a coherent system of analysis and prediction, its practitioners should be able to apply the system in a reliable and convergent manner. In other words, even though the predictions by the six astrologers in this study were incorrect, these predictions still should show a pattern of internal consistency or interastrologer agreement. Pairwise comparisons between astrologers' patterns of predictions for the 23 test cases yielded a mean of only 1.4 agreements. Across the 15 pairwise comparisons among the six astrologers, the number of agreements ranged from zero to three. These results are not significantly better than we would expect by chance. Thus, the astrologers failed to demonstrate interjudge reliability or convergence. Each astrologer apparently was employing an idiosyncratic system to arrive at predictions.

This last finding is particularly troubling and instructive. It implies that, on average, each horoscope could be confidently matched to (at least) six different individuals. We believe that this result was almost inevitable given the nature of the horoscope. The horoscope is extremely complex, providing hundreds of often contradictory "predictions" about an individual. This overcomplexity requires the astrologer to emphasize certain aspects of the chart and downplay others. The many possible combinations that result from assigning different weightings to various elements of the chart very likely produces the result obtained. Indeed, one may be able to find confirmation in the chart for nearly anything one might want to find, at least from the aspect of personality characteristics. Aspects of timing are probably less overdetermined.

One final point should be mentioned. The experimental task probably was considerably easier and, presumably, easier to perform accurately, than the task that astrologers attempt in their counseling practices. That is, without a priori information, because each individual is unique, in practice an astrolo-

ger must use the birth information to "select" the one correct interpretation that uniquely matches that individual from nearly countless possibilities, not just from 23 possibilities. Thus, our task can be seen as a simplification of the task that astrologers routinely undertake as a part of their daily professional practice. The conclusion one can draw from this inference is unequivocal. If our task provides a simplification of standard astrological practice, and if the astrologers cannot perform a simplified task accurately, then it is not likely that they will be able to perform a more complicated task accurately.

Summary

The purpose of this study was to test the central claim of astrology: namely, that conscientious and qualified astrologers can make valid predictions of personal characteristics and life events based solely on knowledge of the date, time, and place of an individual's birth; to ensure that the study would be a fair and representative test of astrology, it was designed and conducted with the full participation and approval of the Indiana Federation of Astrologers. This unique collaboration provided a rare scientific opportunity for an open-minded, critical test.

The results were clear-cut. Six expert astrologers failed to do significantly better than chance or than a nonastrologer control subject at matching birth information to the corresponding case materials for 23 individuals. The astrologers and control subject also did no better at the matching task than ten judges who attempted to rank order the ages of the 23 test cases solely on the basis of photographs. Astrologers' predictive accuracy was unrelated to their level of confidence in their predictions. Furthermore, there was little or no predictive agreement among the astrologers, even though the astrologers purported to be using the same system and methods to arrive at their predictions. Overall, the astrologers probably could have done just as well if they had matched the birth information with the case materials in a random manner.

The first author was required to present the results of the study to the entire IFA membership as part of the agreement we negotiated. During that meeting, the astrologers in attendance broke into groups to attempt the matching task on subsets of five horoscope-personal data datasets—they failed.

These negative results surprised the members of the Indiana Federation of Astrologers, who had remained confident in the predictive powers of astrology throughout the study. The corresponding secretary of the IFA published a report of the project in the *Journal of Research of the American Federation of Astrologers* under the title "Encounter with Academia" (the first author's name was changed to Walter McIntire) and made the following observations about what went wrong:

. . . in many cases, the correct answer contained the attributes we had chosen, but in a different [astrological] position. . . one big mistake was in agreeing to use young

subjects. This was the **Saturn/Neptune** conjunction group, of course, which produced many 'lost souls' . . . Like medicine, the law, and theology, astrology may not always give quantifiable results — but it works, nonetheless. (Mull, 1986)

This response to the study raises interesting questions about the nature of belief systems and the resistance of belief systems to change in the face of disconfirming evidence (Tversky, & Kahneman, 1974).

References

- Bok, B. J., Jerome, L. E., & Kurtz, P. (1975, September). Objections to astrology: A statement by 186 leading scientists. *The Humanist*, 4-6.
- Carlson, S. (1985). A double-blind test of astrology. *Nature*, 318, 419-425.
- Clark, V. (1961, Winter/Spring). Experimental astrology. *In Search*, 102-112.
- Clark, V. (1970, October). An investigation of the validity and reliability of the astrological technique. *Aquarian Agent*, 1, 2-3.
- Dobyns, Z. P. (1976, April). Results of the Vernon Clark experiment. *Astrology Now*, 1, 65, 81.
- Eysenck, H. J., & Nais, D. K. B. (1982). *Astrology, science or superstition?* New York: St. Martin's Press.
- Feller, W. (1961). *An introduction to probability theory and its application*. New York: Wiley Books.
- Gauquelin, M. (1973). *The cosmic clocks*. London: Paladin.
- Gauquelin, M. (1978). *Cosmic influences on human behavior*. New York: ASI.
- Joseph, R. A. (1975). A Vernon Clark model experiment distinguishing exceptionally gifted high performance from profoundly retarded low performance children. *Journal of Geocosmic Research*, 1, 55-72.
- Mayo, J., White, O., & Eysenck, H. J. (1978). An empirical study of the relation between astrological factors and personality. *Journal of Social Psychology*, 105, 229-236.
- Mull, C. S. (1986). Encounter with Academia. *Journal of Research of the American Federation of Astrologers*, 3, 49-52.
- Rockwell, T., Rockwell, R., & Rockwell, W. T. (1978). Irrational rationalists: A critique of *The Humanist's* crusade against parapsychology. *Journal of the American Society for Psychical Research*, 72, 23-34.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185, 1124-1131.
- Vidmar, J. E. (1979, March). Astrological discrimination between authentic and spurious birth-dates. *Cosmecology Bulletin*, 8 & 9, 57-90.