

Support for Astrology from the Carlson Double-blind Experiment

Ken McRitchie

This article has been peer reviewed by subject matter experts refereed through the publisher.

ISAR International Astrologer, Volume 40, Issue 2, August, 2011, Pages 34-39

(submitted October 24, 2009)

Abstract

The Carlson double-blind study, published in 1985 in Nature (one of the world's leading scientific publications) has long been regarded as one of the most definitive indictments against astrology. Although this study might appear to be fair to uncritical readers, it contains serious flaws, which when they are known, cast a very different light on the study. These flaws include: no disclosure of similar scientific studies, unfairly skewed design, disregard for its own stated criteria of evaluation, irrelevant groupings of data, rejection of unexpected results, and an illogical conclusion based on the null hypothesis. Yet, when the stated measurement criteria are applied and the data is evaluated according to normal social science, the two tests performed by the participating astrologers provide evidence that is consistent with astrology ($p = .054$ with $ES = .15$, and $p = .037$ with $ES = .10$). These extraordinary results give further testimony to the power of data ranking and rating methods, which have been successfully used in previous astrological experiments. A critical discussion on follow-up studies by McGrew and McFall (1990), Nanninga (1996/97), and Wyman and Vyse (2008) is also included.

The research experiment conducted by Shawn Carlson, "A double blind test of astrology," published in the science journal *Nature* in 1985 as an indictment of astrology, is one of the most frequently cited scientific studies to have claimed to refute astrology. A Google search for the title as a quoted string returns over 6,600 links.¹ Although the Carlson study drew initial criticism for numerous flaws when it was published, a more recent examination has found that despite the flaws, the data from the study actually supports the claims of the participating astrologers. This support lends further credence to the effectiveness of ranking and rating methods, which have been used in other, lesser known astrological experiments.

The Carlson astrology experiment was conducted between 1981 and 1983 when Carlson was an undergraduate physics student at the University of California at Berkley under the mentorship of Professor Richard Muller. The flaws that have

been uncovered in the *Nature* article include not only the omission of literature on similar studies, which is expected in all academic papers, but more serious irregularities such as skewed test design, disregard for its own criteria of evaluation, irrelevant groupings of data, removal of unexpected results, and an illogical conclusion based on the null hypothesis.

In concept and design, the Carlson experiment was not original. It was modeled after the landmark double-blind matching test of astrology by Vernon Clark (Clark, 1961). In that test astrologers were asked to distinguish between each of ten pairs of natal charts. One chart of each pair belonged to a subject with cerebral palsy and the other belonged to a subject with high intelligence. Another influential study was the "Profile Self-selection" double-blind experiment, which was led by the late astrologer Neil Marbell and privately distributed among contributors in 1981 before its eventual publication (Marbell, 1986-87). In that test, participating volunteers were asked to select their own personality interpretations, both long and short versions in separate tests, out of three that were presented.

In both of these prior studies, the participants performed well above significance in support of the astrological hypothesis as compared to chance. The Marbell study was extraordinarily qualified as it involved extensive input and review from astrologers, scientists, statisticians, and prominent skeptics. Carlson neglected to provide any review of these scientific studies that supported astrology or any other previous related experiments.

The stated purpose of Carlson's research was to scientifically determine whether the participating astrologers (members of the astrology research organization NCGR and others) could match natal charts to California Psychological Inventory (CPI) profiles (18 personality scales generated from 480 questionnaire items). Additionally, Carlson would determine whether participating volunteers (undergraduate and graduate students, and others) could match astrological interpretations, written by the participating astrologers, to themselves. These assessments, Carlson asserts, would test the "fundamental thesis of astrology" (Carlson, 1985: 419).

From the time of its release, the Carlson study has been criticized for the extraordinary demands it placed on the

participating astrologers, which would be regarded as unfair in normal social science. As with any controversial study, all references to Carlson's experiment should include the scientific discourse that followed it, particularly the points of criticism that show weaknesses in the design and analysis. Notable among recent critics has been University of Göttingen emeritus professor of psychology Suitbert Ertel, who is an expert in statistical methods and is known for his criticism of research on both sides of the astrological divide. Ertel published a detailed review in a 2009 article, "Appraisal of Shawn Carlson's Renowned Astrology Tests" (Ertel, 2009).

From a careful reading of Carlson's article in light of the ensuing body of discourse, we can appreciate that the design of the experiment was intentionally skewed in favor of the null hypothesis (no astrological effect), which Carlson refers to, somewhat misleadingly as the "scientific hypothesis." Some of the controversial features of the design are as follows:

- The astrologers were not supplied with the gender identities of the CPI owners, even though the CPI creates different profiles for men and women. (Eysenck, 1986: 8; Hamilton, 1986: 10).
- Participants were not provided with sufficiently dissimilar choices of interpretations, as the Vernon Clark study had done, but instead were given randomly selected choices. This may give the impression of a fair method, but given the narrow demographics of the sample, there is an elevated likelihood of receiving similar items from which to choose, which makes it unfair (Hamilton, 1986: 12; Ertel, 2009: 128).
- The easier to discriminate and more powerful two-choice format, which had been used in the Vernon Clark study, was replaced with a less powerful three-choice format, which further elevated the chances of receiving similar items (Ertel, 2009: 128). No reasons are given for this unconventional format, although it can be surmised that Carlson was well aware of the complexities of a three-choice format from his familiarity with the Three-Card Monte ("Follow the Lady") sleight of hand confidence game, which he had often played as a street psychic and magician (Vidmar, 2008).
- The requirement for rejecting the "scientific hypothesis" was elevated to 2.5 standard deviations above chance ($p = .006$). In the social sciences, the conventional threshold of significance is 1.64 standard deviations with probability less than $p = .05$ (Ertel, 2009: 135).
- Failure to consider the astrologers' methodological suggestions or give an account of their objections. Carlson credits astrologer Teresa Hamilton with giving "valuable suggestions," yet Hamilton complained later that "Carlson followed none of my suggestions. I was never satisfied that

the experiment was a fair test of astrology" (Hamilton, 1986: 9).

Given this skewed design, the irregularities of which are not obvious to the casual reader, Carlson directs our attention to the various safeguards he used to assure us that no unintended bias would influence the experiment. He describes in detail the precautions used to screen volunteers against negative views of astrology, how the samples were carefully numbered and guarded to ensure they were blind, and the contents of the sealed envelopes provided to test participants.

The experiment consisted of several separate tests. The astrologers performed two tests, a CPI ranking test and a CPI rating test. The volunteer students performed three tests, a natal chart interpretation ranking test, a natal chart interpretation component rating test, and a CPI ranking test.

In the CPI ranking test, astrologers were given, for each single natal chart, three CPI profiles, one of which was genuine, and asked to make first and second choices. There were 28 participating astrologers who matched 116 natal charts with CPIs. Success, Carlson states, would be evaluated by the frequency of combined first and second choices, which is the correct protocol for this unconventional format. He states, "Before the data had been analyzed, 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" (Carlson, 1984: 425).

In addition to this ranking test, the astrologers were tested for their ability to rate the same CPIs according to a scale of accuracy. This task allowed for finer discrimination within a greater range of choices. Each astrologer "also rated each CPI on a 1-10 scale (10 being the highest) as to how closely its description of the subject's personality matched the personality description derived from the natal chart" (Carlson, 1985: 420).

As to the results of the astrologers' three-choice ranking test, Carlson first directs our attention to the frequency of the individual first, second, and third CPI choices made by the astrologers, each of which he found to be consistent with chance within a specified confidence interval. This observation is scarcely relevant, given the stated success criteria of the first and second choice frequencies combined. Then, to determine whether the astrologers were successful, Carlson directs our attention to the rate for the third place choices, which, as already noted, was consistent with chance. Thus he declares that the combined first two choices were *not* chosen at a significant frequency.

"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 [sic] the correct CPI as their first or second choices at a significant

level” (Carlson, 1984: 425). This conclusion, however, ignores the stated success criteria and is in fact untrue. The calculation for significance shows that the combined first two choices were chosen at a success rate that is marginally significant ($p = .054$) (Ertel, 2009: 129).

As to the results of the astrologers’ rating test (10-point rating of three CPIs against each chart), Carlson demonstrates that the astrologers’ ratings were no better than chance within the first, second, and third place choices made in the three-choice test. He shows a weighted histogram and a best linear fit graph to illustrate each of these three groups of ratings. Carlson directs our attention to the first choice graph as support for his conclusion for this test. The slope of this graph is “consistent with the scientific prediction of zero slope” (Carlson, 1985: 424). The slope is actually slightly downward. The graphs for the other two choices are not remarked upon, but show slightly positive slopes.

The notable problem with Carlson’s analysis of the 10-point rating test, however, is that this test had no dependency on the three-choice ranking test and even used a different sample size of CPIs.² According to the written instructions supplied to the astrologers, this rating test was actually to be performed *before* the three-choice ranking test (Ertel, 2009: 135). These 10-point ratings should not be grouped as though they were quantitatively related to the later three-choice test. Confirmation bias from the claimed “result” of the three-choice test, which Carlson presents earlier in his paper, suggests acceptance of irrelevant groupings in this 10-point rating test, presented later. When the totals of the ratings are considered without reference to the choices made in the subsequent test, a positive slope is seen, which shows that the astrologers actually performed at an even higher level of significance ($p = .037$) than the three-choice test (Ertel, 2009: 131).

The other part of Carlson’s experiment tested 83 student volunteers to see if they could correctly choose their own natal chart interpretations written by the astrologers. Volunteers were divided into a test group and a control group. Members of the test group were each given three choices, all of the same Sun sign, one of which was interpreted from their natal chart (Carlson, 1985: 421). Similarly, each member of the control group received three choices, all of the same Sun Sign, except none of the choices was interpreted from their natal charts, although one choice was randomly selected as “correct” for the purpose of the test.

For the results of this test, Carlson shows a comparison of the frequencies of the correct chart as first, second, and third choices for the test group and the control group (again ignoring his stated protocol to combine the frequencies of the first two choices). He finds that the results for the test group are “all consistent with the scientific hypothesis” (Carlson, 1985: 424). However, he does note an unexpected

result for the control group, which was able to choose the correct chart at a very high frequency. He calculates this to be at 2.34 standard deviations above chance ($p = .01$). Yet, because this result occurred in the control group, which was not given their own interpretations, Carlson interprets this as a “statistical fluctuation.”

Yet the size of this statistical fluctuation is so unusual as to attract skepticism, particularly in light of Carlson’s other results. It is reasonable to think that the astrologers could write good quality chart interpretations after having successfully matched charts with CPI profiles. Yet, according to Carlson’s classification, the test group tended to avoid the astrologers’ correct interpretations and choose the two random interpretations, while the control group tended to choose the selected “correct” interpretations by a wide margin, as if they, the controls, had been the actual test subjects (Ertel, 2009: 132). This raises suspicion that the data might have been switched, perhaps inadvertently, but this is unverifiable speculation (Vidmar, 2008).

Like the participating astrologers, the student volunteers were also given a rating test; in this case for the sample chart interpretations they were given. They were asked to rate, on a scale of 1 to 10, the accuracy of each subsection of the natal chart interpretations written by the astrologers. “The specific categories which astrologers were required to address were: (1) personality/temperment [sic]; (2) relationships; (3) education; (4) career/goals; and (5) current situation” (Carlson, 1985: 422). This test would potentially have high interest to astrologers because of the distinction it made between personality and current situation, which is a distinction that is not typically covered in personality tests. Also, the higher sensitivity of a rating test could provide insight, at least as confirmation or denial, into the extraordinary statistical fluctuation seen in the three-choice ranking test.

However, based on a few unexpected results, Carlson decided that there was no guarantee that the participants had followed his instructions for this test. “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” (Carlson, 1985: 424). On the basis of this unanticipated consistency, Carlson rejected the volunteers’ rating test without reporting the results.

As an additional test in this part of the experiment, the student volunteers were asked to choose from among three CPI profiles the one that was based on the results of their completed CPI questionnaire. The other two profiles offered were taken from other student volunteers and randomly added. Of the 83 volunteers who completed the natal chart interpretation choices, only 56 completed this task. As usual, Carlson compared the results of the three choices for the test and control groups taken individually (instead of the

frequency of the first two choices taken together). Furthermore, in contravention to the logic of control group design, Carlson compares the two groups against chance instead of against each other (Ertel, 2009: 132). He found no significant difference from chance for the two groups.

There are plausible reasons that could explain why the test group was unable to correctly select their own CPI profiles, even though the astrologers were able to a significant extent as we have seen, to match CPI profiles with the students' charts. The disappointing number of students who completed this task, despite having endured the 480-question CPI questionnaire, suggests that the students might have been much less motivated than the astrologers, for whom the stakes were higher (Ertel, 2009: 133). The CPI matching tasks, for both the volunteers and the astrologers, were especially challenging because of the three-choice format. The random selections of CPIs made within the narrow demographics of the sample population of students would have elevated the likelihood of receiving at least two CPI profiles that were too similar to make a discriminating choice and this would have had a negative impact on motivation.

In the conclusion of his study, Carlson claims: "We are now in a position to argue a surprisingly strong case against astrology as practiced by reputable astrologers" (Carlson, 1985: 425). However, this conclusion defies rationality. Ertel points out the logical flaw that such a conclusion cannot be drawn even if the tests had shown an *insignificant* result. "Not being able to reject a null hypothesis does not justify the claim that the alternate hypothesis is wrong" (Ertel, 2009: 134).

Despite its numerous flaws and unfair challenges, the Carlson experiment nevertheless demonstrates that the astrologers, in their two tests, were able to match natal charts with CPI profiles significantly better than chance according to the criteria normally accepted by the social sciences. Thus the null hypothesis must be rejected. As such, the Carlson experiment demonstrates the power of ranking and rating methods to detect astrological effects, and indeed helps to raise the bar for effect size in astrological studies. The benchmark effect size that had been attained by the late astrological researcher Michel Gauquelin was merely .03 to .07. Although these were small effects, they were statistically very significant due to large sample sizes (N = 500-1000 or more natal data) and had to be taken seriously (Gauquelin, 1988a). In Carlson's experiment, which applied sensitive ranking controls, the effect size of the three-choice matching test with $p = .054$ is $ES = .15$, and the effect size of the 10-point rating test with $p = .037$ is $ES = .10$ (Ertel, 2009: 134).

Follow-up studies

Other experiments have attempted to address the earlier documented criticisms of the Carlson test. However, these experiments, each of which claims to confirm that astrological choices are made at no better than chance levels, have drawn

criticism from astrologer Robert Currey (2011) and others as having fatal flaws. Each falls short of the Carlson study. Included here are the studies by McGrew and McFall (1990), Nanninga (1996/97), and Wyman and Vyse (2008).

The McGrew and McFall (1990) experiment was intended to include personal information of the sort typically used by astrologers but not found in standard personality profiles. Six "expert" astrologers, all members of the Indiana Federation of Astrologers but none of whom claimed professional accreditation, participated. Each astrologer was asked to match the birth charts of 23 volunteers to an extremely broad range of information gathered for each volunteer. This information included photo portraits, results from two standardized psychology tests, and written descriptions of personality and life events generated by 61 questions that were developed from input that the authors gleaned from the astrologers.

The use of photos in the McGrew and McFall study meant that special restrictions were imposed on the experiment to avoid age clues from the photos. The authors recruited volunteers who ranged from only 30 to 31 years of age. This narrow demographic, where natal charts would share numerous similarities, and the large amount of non-uniform information supplied for each volunteer, elevated the difficulty of the matching task. The Carlson study is regarded as unnecessarily complex because the astrologers were asked to choose the genuine CPI from among three. In the McGrew and McFall study however, astrologers were given the virtually impossible task of choosing each genuine set of personal descriptions and information from among no less than 23 sets! The authors argue that the astrologers' experimental task was a "simplification" of their ordinary business (McGrew and McFall, 1990: 82). On the contrary, it was much more complex and far more difficult than even Carlson's tasks. The reasons that the two authors provide for their judgment against astrology is not at all convincing.

The Nanninga (1996/97) experiment was modeled on the McGrew and McFall experiment and contained the same sorts of flaws. It was intended to settle a dispute argued in the local newspapers as to whether astrologers can or cannot predict. Through the newspapers, Nanninga offered a large cash prize to anyone who could match seven natal charts to seven sets of personality information. He attracted an unexpectedly large number of "astrologers," from which he chose 50 based on their claimed astrological experience. The test subjects for the study were volunteers, all born "around 1958." A test questionnaire for the volunteers, developed by Nanninga from ideas solicited from the astrologers, covered a very wide range of interests and background such as education, vocation, hobbies, interests, main goals, personality, relationships, health, religion, and so on, plus dates of important life events. To these Nanninga added 24 multiple choice questions taken from a standard personality test.

Like the McGrew and McFall experiment, Nanninga's experiment used a very narrow demographic of volunteer subjects, making them difficult to astrologically differentiate, and he likewise presented a very large amount of non-uniform personal data written by the seven volunteers for the astrologers to sort through. Although Nanninga's task involved seven matches instead of 23 and was therefore somewhat less complex than the McGrew and McFall task, it was nonetheless considerably more complex than the Carlson task, which has been criticized as being more complex than necessary. Nanninga's study was not an improvement over the Carlson experiment and does not convincingly support his claims that astrology is in conflict with science and that astrologers increasingly confine themselves to statements that cannot be falsified (Nanninga, 1996/97: 20).

The Wyman and Vyse (2009) experiment was a low-budget classroom study modeled on the Carlson experiment but without the astrologers. In this experiment it was hypothesized that the use of a very transparent self-assessment questionnaire (the NEO Five-Factor Inventory) would enable volunteer participants to better identify their own profile scores than the CPI used by Carlson. Examples from this questionnaire include, "I try to be courteous to everyone I meet" (which contributes to A, Agreeableness in the resultant profile), and "I like to be where the action is" (which contributes to E, Extraversion). The authors asked 52 volunteers (introductory psychology class members and others) to identify their genuine five-factor personality profile from a bogus one and to identify their genuine astrological description from a bogus one. The astrological descriptions were created from the output of a commercial natal chart interpretation program, modified to remove all planetary, sign, and house clues and further simplified by the removal of all aspect information to provide 29 one- to four-sentence personality descriptions. The students succeeded at the personality profile task but failed at the natal chart description task.

Criticisms of the Wyman and Vyse experiment include: (1) No test of astrologers' skills and performance. (2) The false assumption that both natal chart interpretations and psychology profiles "share a common purpose — to provide a description of the respondent's personality" (Wyman and Vyse, 2008: 287). Natal charts provide their value as descriptions of potential. (3) The tender age of the volunteers (mean age of 19.3 years) whose life potential would be largely unrealized and somewhat idealized. (4) Small sample size of natal charts (N = 52, where a sample of 100 would have been better). (5) The exclusion of aspects from the astrological descriptions, arguably the most important component. (6) Lack of synthesis of the chart components and a holistic approach. (7) The unbalanced tasks of identifying an easy five-factor profile that parrots the subject's input compared to the complexity of identifying a 29-factor partial astrological description of life potential. (8) The false assumption that the

positive and negative polarities of the signs mean "favorable" and "unfavorable" respectively and the listing (twice) of the sign Aquarius as both favorable and unfavorable. (9) Incomplete disclosure of result details. Statistical inferences were drawn based on belief in astrology, but how many students in this small sample would dare, even anonymously, to declare belief in astrology in an experiment presided over by a professor, Stuart Vyse, who is a prominent astrology skeptic? Was it more than one? (10) Students' fear for their academic safety is a high stakes issue and could easily bias such as study as this one.

These errors and inadequacies arouse suspicion as to the accuracy of the modified astrological descriptions. Together, these flaws place the Wyman and Vyse experiment well below the level of the Carlson experiment and raise serious doubts as to the authors' conclusions. Although the simple five-factor personality profiles were identifiable by the students at a significant rate, the authors' claim that the simplified astrological descriptions they devised should be equally identifiable is not convincing.

Discussion

The evidence provided by the Carlson experiment, when considered together with the scientific discourse that followed its publication, is extraordinary. Given the unfairly skewed experimental design, it is extraordinary that the participating astrologers managed to provide significant results. Given the irregularities of method and analysis, which had somehow remained transparent for 25 years, it is extraordinary that investigators have managed to scientifically assess the evidence and bring it into the full light of day. Now that the irregularities have been pointed out, it is easy to see and appreciate what Carlson actually found.

However, because of the unfairness and flaws in the Carlson experiment, this line of research needs to be replicated and extended in more stringent research programs. The research done in the follow-up studies by McGrew and McFall (1990), Nanninga (1996/97), and Wyman and Vyse (2008) were on the whole better executed with regard to method and analysis than the Carlson experiment. However, first-rate execution does not magically transform faulty assumptions and design into first-rate science. These are the relatively easy and routine parts of research that can often be rescued from their own problems, as we have seen with the Carlson study. With hindsight, it is evident that the editors of the science and psychology journals who published these research studies failed to realize that astrology is a complex discipline with many variables, limitations, and pitfalls. Ultimately, it is important that astrologers offer criticism back to their critics and help them to avoid the fundamental blunders and misjudgments outlined in this article. Astrological expertise should always be included in the final peer review stage prior to publication.

There is much to be learned from the Carlson experiment. If natal charts can be successfully compared with self-assessment tests by the use of rating and ranking methods, as the Carlson experiment indicates, then astrological features might be easier to evaluate than was previously believed. New questions must now be raised. What would the results be in a fair test? Why did the astrologers choose and rate the CPIs as they did? Which chart features should be compared against which CPI features? Could more focused personality tests provide sharper insights and analysis? The door between astrology and psychology has been opened by a just crack and we have caught a glimpse of hitherto unknown connections between the two disciplines.

Notes

1. By comparison, a Google query of some other peer reviewed journal articles on astrology, searched as quoted strings, returns the following: "Is Astrology Relevant to Consciousness and Psi?" (Dean and Kelly, 2003) 8800 results; "Are Investors Moonstruck?—Lunar Phases and Stock Returns" (Yuan et al, 2006) 3700 results; "Objections to Astrology: A Statement by 186 Leading Scientists" (*The Humanist*, 1975) 3500 results; "A Scientific Inquiry Into the Validity of Astrology" (McGrew and McFall, 1990) 2160 results; "Raising the Hurdle for the Athletes' Mars Effect" (Ertel, 1988) 1350 results; "The Astrotest" (Nanninga, 1996) 970 results; "Is There Really a Mars Effect?" (Gauquelin, 1988) 630 results; "Science versus the Stars" (Wyman and Vyse, 2008) 420 results.
2. Carlson presents the 10-point rating test as a finer discrimination of the 3-choice ranking test, but the sample size is not the same. A sample of 116 natal charts is used in the 3-choice test (Carlson, 1985: 421, 423) and a different sample size is used for the 10-point rating test, which adds to the discrepancies already mentioned between these two tests and further emphasizes that they cannot be considered as a single test. Carlson does not give the sample size for the 10-point test, but it can be determined by measurement of the first, second, and third choice histograms in his article (Carlson, 1985: 421, 424). Each natal chart had to be the "correct" choice in one of these three "choices." By adding up these "correct hits," Ertel shows 99 charts (Ertel, 2009: 130, Table 3). A more exacting scrutiny of the histograms by Robert Currey (in a forthcoming article) determines 100 charts.

References

Carlson, Shawn (1985). "A double-blind test of astrology," *Nature*, 318, 419-425.

Clark, Vernon (1961). "Experimental astrology," *In Search*, Winter/Spring, 102-112.

Currey, Robert (2011). "Research Sceptical of Astrology: McGrew & McFall, 'A Scientific Inquiry into the Validity of Astrology' 1990." Retrieved on 2011-07-02.

Currey, Robert (2011). "Research Sceptical of Astrology: Wyman & Vyse Double Blind Test of Astrology." Retrieved on 2011-07-02.

Ertel, Suitbert (1988). "Raising the Hurdle for the Athletes' Mars Effect: Association Co-varies with Eminence" *Journal of Scientific Exploration*, 2(1), 53-82.

Ertel, Suitbert (2009). "Appraisal of Shawn Carlson's Renowned Astrology Tests," *Journal of Scientific Exploration*, 23(2), 125-137.

Eysenck, H.J. (1986). "A critique of 'A double-blind test of astrology'" *Astropsychological Problems*, 1(1), 27-29.

Gauquelin, Michel (1988). "Is there Really a Mars Effect?" *Above & Below: Journal of Astrological Studies*, Fall, 4-7.

Hamilton, Teresa (1986). "Critique of the Carlson study" *Astropsychological Problems*, 3, 9-12.

Marbell, Neil (1986-87). "Profile Self-selection: A Test of Astrological Effect on Human Personality," *NCGR Journal*, Winter, 29-44.

McGrew, John H. and Richard M. McFall (1990). "A Scientific Inquiry Into the Validity of Astrology," *Journal of Scientific Exploration*, 4(1), 75-83.

Nanninga, Rob. (1996/97). "The Astrotest: A tough match for astrologers." *Correlation*, Northern Winter, 15(2), 14-20.

Vidmar, Joseph (2008). "A Comprehensive Review of the Carlson Astrology Experiments." Retrieved on 2010-08-01.

Wyman, Alyssa Jayne and Stuart Vyse (2008). "Science Versus the Stars: A Double-Blind Test of the Validity of the NEO Five-Factor Inventory and Computer-Generated Astrological Natal Charts." *The Journal of Psychology*, 135(3), 287-300.

Biography



Ken McRitchie is a Canadian poet, technical writer, and research astrologer. He is the one-time editor (1985-89) of *Above & Below: Journal of Astrological Studies*, and is the author of *Environmental Cosmology: Principles and Theory of Natal Astrology*. His website is www.theoryofastrology.com.

Email ken@theoryofastrology.com.