U- turn in Carlson's astrology test?

by

Robert Currey

Introduction

New Scientific Support for astrology comes from a surprising source.

In 1985, Nature, one of the world's most prestigious scientific journals, published "A Double Blind Test of Astrology", now known as the 'Carlson experiment'. The conclusion was that natal astrology as practised by reputable astrologers was no better than chance. For astrology, it was a landmark experiment that continues to undermine the credibility of every astrological consultant, researcher and school. Recent research shows not only that Carlson's conclusions were wrong (Vidmar 2008), but also that his experiment produced evidence that the tested astrologers performed their tasks successfully to a level that could not be explained by chance (Ertel 2009).

This article attempts to synthesise the evidence from different sources with some additional observations. It seeks to clarify how Carlson imported the results of one test into another separate test. This led to sampling errors that disguised results that favoured the astrologers - results that were later discovered by Professor Ertel. To balance this, an attempt is made, here, to present the sceptical reaction to this new evidence. Graphics have been compiled to display the astrologer's predicted rating of their matches weighted by the frequency and to show how enabling astrologers to make these confidence judgements with the data amplifies the precision.

Background

In March 2010, the BBC broadcast a spectacular astronomy series entitled the 'Wonders of the Solar System' in the UK. Twenty-six minutes into series 4, the presenter physicist, Professor Brian Cox, made the following remark about astrology:

"*Astrologists have said for years that Jupiter influences our lives. But we now have scientific evidence that this mighty planet does have a significant connection with our own small world.*

Now, Jupiter is so different to [sic] our planet… a big ball of gas half a billion kilometres away. It's difficult to see how it could have anything to do with us at all. But ***despite the fact that astrology is a load of rubbish****, Jupiter can in fact, have a profound influence on our planet. And it's through a force … gravity."*

It was an unprofessional remark from someone who knew nothing about astrology. Cox allowed his bias to intrude into an educational scientific programme broadcast the BBC, the guidelines of which state that astrology must be broadcast in a balanced way. [1]

Enquiries about this article should be addressed to the author at: **robert@equinoxastrology.com**

Normally this means astrology may be presented with confrontational critics, but it should also entitle astrologers to defend their field. After I expressed my views on the social networks, Cox responded by 'retweeting' my comments on Twitter. Overnight, Phil Plait, former President of the James Randi Educational Foundation and self-styled "Bad Astronomer," wrote a blog (Plait 2010) about my article on his sceptical website: badastronomy.com. Many hundreds of hard-core sceptics clicked the link to my Facebook public page (Facebook 2010) to offer a mixture of insult, ridicule and to tell me why astrology really is rubbish.

When I responded rationally, a few interesting debates and even dialogue followed. Like most astrologers, I value subjective evidence in the form of personal experience and observation, but I also value objective tests. And, in my view, while the practice of astrology by most astrologers is best described as an art or a craft, there is no scientific case for dismissing astrology as a field.

The BBC replied to my complaint with an apology and a reassurance that the Professor's comments, *"were his own, not those of the BBC and were based on his belief that there isn't sufficient evidence to support astrology"*.

The Myth of Scientific Evidence against Astrology

How did we ever get to this? How can educated scientists know so little about astrology? What makes an eminent Professor and particle physicist at the Large Hadron Collider on the French-Swiss border decide to bypass the scientific imperative of empirical study and broadcast unfounded conclusions?

The answer came back more quickly than I had thought. Hordes of sceptics asserted Cox's belief that there is indeed no scientific evidence to support astrology and that there is a preponderance of tests that shows that astrology is no better than a sophisticated game of dice.

When challenged, the sceptics cited the work of Dr Geoffrey Dean. Dean, an analytical chemist by training and, once an astrologer, is arguably the most vocal, informed and famous critic of astrology. They cited his Meta-Analysis of 300 experiments. (Dean 1997). However, the certainty of the conclusions is not supported by the References, where fourteen years later it still states, "*A list of the studies meta-analysed above is in preparation*".

Dean's impressive sounding Time-Twin study involving 2,101 people born in London between 3rd and 9th May 1958 was also cited. This sounded promising, as Dean gives a trailer in his paper on Astrology and PSI (Dean & Kelly 2003) (Dean, forthcoming). Yet, to date (eight years on), Dean has still not published his results.

Since their belief could not be supported by science, several sceptics cited what may be perceived as conjuring tricks allegedly debunking astrology. In fact, there are sound

reasons why pseudo-astrological illusions by magicians such as James Randi (Currey July 2010) or Derren Brown (Currey Aug. 2010) cannot be reproduced under scientific conditions.

However, sceptics and sceptical websites repeatedly cite one solid test above all others. It is the famous [2] Carlson Double-Blind Astrology test (Carlson 1985) that was published in *Nature* in 1985. Pearce Wright, science editor of *The Times* dedicated an unprecedented 5 page article to the experiment to coincide with publication. Shawn Carlson's mentor, Professor Richard Muller describes it as the "definitive test of astrology". (Muller 2010) It is cited in Wikipedia under Astrology along with Dean's "phantom" Time-Twin Test (where the results still appear to be unpublished) with the claim: *"Studies have repeatedly failed to demonstrate statistically significant relationships between astrological predictions and operationally defined outcomes*." (Wikipedia 2010) While Ptolemy or natural astrology is currently not mentioned on the main Astrology listing, a few editors zealously protect numerous spurious claims. In March 2011, an administrator banned six of the most active editors with the required astrological expertise from editing astrology pages on the grounds of an in-house technical loophole (Wikipedia 2011).

Carlson Double-Blind Astrology Experiment

Though it is now widely known as the "Carlson Test", the experiment bears the hallmark of Committee for Scientific Investigation of Claims of the Paranormal also known as CSICOP [3]. This group, since renamed Committee for Skeptical Inquiry (CSI), is arguably dedicated to debunking metaphysical evidence. In the early 80s CSICOP was in crisis and many of its senior members resigned in protest over the "Gauquelin affair" [4]. In short, their highly questionable attempt to refute Michel Gauquelin's evidence supporting astrology failed. What was worse for CSICOP was that their own data replicated Gauquelin's results. (Ertel & Irving 1996) (Rawlins 1981) Following this, CSICOP abandoned running in-house experiments. Nevertheless with Carlson, CSICOP was involved in encouraging, funding, guiding and publicising the experiment [5].

Shawn Carlson was a 19-year-old [6] physics undergraduate student at the University of California, Berkeley when he started this study in 1980. In the years since the experiment was published, Carlson has been active with CSICOP (now CSI) and the James Randi Education Foundation. Like Randi and numerous members of CSICOP, Carlson was also a professional magician and it is hard not to assume that Carlson did not share their sceptical mind-set before the experiment. At the age of 16, he supported himself as a player of Three-Card-Monte (also known as Follow the Lady) [7]. This card trick requires mastery of the art of deception and misdirection.

Did Carlson have a hidden agenda?

Carlson initially gave the participating astrologers the impression that he was favourable towards astrology. During the experiment, he claimed in a signed letter to astrologer Erin Sullivan [undated ca. October 1981] (published in Correlation 26 (1) Vidmar 2008) that the preliminary results appeared to support astrology. *"We are very near interpreting the results as FAVORING the astrological thesis. Near, but not there yet."* After the test was

published, Carlson wrote in a letter to Jayj Jacobs entitled Rebuttal [mailed 21 January 1987]: *"Mr Lewis is half correct on this point. It is true that I was indeed biased. But I wanted the experiment to work."* (Vidmar 2008)

Had the astrologers believed that Carlson had a hidden agenda against astrology when he began this experiment it is likely that they would not have participated. Yet, there is plenty of evidence to suggest that he was not favourably disposed to astrology before and during the experiment. Before he started, Carlson stated his (modest) intention to submit the test for publication in the Skeptical Enquirer. This in-house journal for CSICOP is committed to discrediting exactly the type of astrological evidence that Carlson was supposedly hoping to find. [8] Also, his work was supported by CSICOP members including his mentor Professor Richard Muller, the founder of CSICOP Professor Paul Kurtz and the editor of the science journal *Nature* and CSICOP fellow, Sir John Maddox. (Frazier 2009)

Before going any further, all the points made so far do not invalidate this experiment. However, I will make it clear that this background of apparent prejudice against astrology had a bearing on the methods, presentation, conclusions and publication.

Protocols for Carlson's Experiment as set out in his Paper

A key part of the whole experiment is the measurement of astrological birth charts against the California Psychological Inventory [9] known as the CPI. This 'self-report psychological test' created in 1956, consists of 480 true-false questions resulting in 18 personality attribute scales. Carlson stated that he chose the CPI for this test as his three advising astrologers considered its scales "*were closest to those discernible by astrology*." (Carlson 1985:420)

The Double-Blind Astrology Experiment consisted of five separate, though related, tests. Carlson merged four of them into Parts 1 and 2, which, it is suggested, here, created an illusion of connection enabling him to use one test to bias another.

1. Subject Ranking Test #1: 83 subjects were each given a selection of 3 birth chart interpretations drawn up by the astrologers. The 3 charts consisted of 2 random birth charts belonging to other participants and their own. They were asked to select their first and second choice according to best fit. (Part 1 according to Carlson)

2. Subject Rating Test #2: An undisclosed number of subjects rated each birth chart (out of 10 points) from a selection of three including their own. (Part 1 acc. Carlson)

3. Astrologer Ranking Test #3: The astrologers were given batches of 3 California Psychological Inventories (CPIs) including 1 authentic match and 2 random belonging to other subjects. They were asked for their first and second choice as closest matches with each of 114-116 [10] subject's birth charts. (Part 2 acc. Carlson)

4. Astrologer Rating Test #4: The astrologers rated 308 CPIs (on a scale of 1 to 10) in batches of 3 according to the closeness of each match with 100 subject's birth charts. (Part 2 acc. Carlson)

5. CPI Ranking Test #5: In a non-astrological test, 56 subjects were required to identify and rank their own California Psychological Inventory CPI (completed earlier) from 2 other random CPIs from the subject group.

Twenty-eight astrologers accepted the invitation to participate in the experiment. Carlson does not say how many actually completed their assignment.

Design & Execution Faults that put Astrology at a disadvantage

An apparent fault in Carlson's test design meant that he rejected Test #2 due to the data "*not having been collected under proper controls*." The subjects were asked to rate the interpretation of their own chart on a scale of 1 to 10. However, Carlson noticed from the "*first few data envelopes*" that on "*a subject's first choice, nearly all the subsections were also rated as first choice.*" This is an unfortunate loss. Given the success of the astrologers in rating CPIs, this data may have got around some of the problems inherent in the subject group (outlined below). It would be helpful if Carlson were to release the raw data so we can at least rule out any question of selective reporting and see if some data can be extracted.

Another setback came when the subjects in test #5 were unable to rank their own CPI at a level greater than chance. As a result, Carlson ruled out test #1 where subjects were required to rank their own birth chart and concluded that self-selection was "*a poor test of astrology*". (Carlson 1985)

Carlson imposed abnormally and inappropriately high levels of proof. The convention in social sciences is that P-values (probability) of .05 or less are indicative of statistical significance. [11] With this particular experiment, much larger sample sizes would be needed to demonstrate highly significant P-values. Yet, Carlson demanded a P-value of .006 (Ertel 2009:135) (2.5 standard deviations), more appropriate for physics than a psychology experiment. Carlson admitted that this was unusually high in his article for *Nature*, but his comments were sub-edited before publication - perhaps to make his conclusions more credible. (Vidmar 2008)

Disadvantages that the Astrologers faced in fulfilling their Tasks:

1. Is it a realistic and fair test of astrology?

Astrology is not a guessing game. This *Guess My Sign* or *Find The Lady* test of identification is not what astrologers would normally do for their clients. Everyone involved in the experiment, including the astrologers, seemed to assume that personal data should stand out to the owner from a crowd of similar data. But can students spot their exam results or patients their medical diagnosis without their name at the top? Blind selection proved challenging for subjects picking their own psychological profile (CPI) [Test #5] or for selecting their birth chart [Test #1]. The astrologers

were, however, more successful in comparing birth charts with the matching psychological profile to a significant level [Tests #3 & #4].

2. Test Subjects: too inexperienced, too homogenous and too disinterested

Astrologers were asked to match a birth chart to one of three CPI profiles. However, the subjects had too much in common with each other for easy identification and separation. The majority (70%) of the subjects were, like Carlson, students at the same University. This high level of homogeneity makes it hard to differentiate between individuals. In addition, the subjects were not predisposed to self-enquiry [12] and their self-knowledge was low. Being an average age of 28 [13] suggests that most would not have undergone their formative first Saturn Return (29-30) and be less able to identify their future potential and inner nature according to astrological theory and conventional wisdom. Most would not have encountered the character-forming experiences of a career, marriage, children and home-ownership for example. They were as Vidmar puts it "*heavily influenced by socio-cultural factors such as parents ...*" (Vidmar 2008) Test-retest results of the CPI on other subjects suggest that after 25 years, half the individuals had altered their self-image by more than 40%.[14] Finally, Carlson's subjects were half-hearted in their attempts to complete their psychological questionnaire and may have been prone to response bias. [12] These limitations were disadvantageous both to astrology and psychology in all the tests.

3. Limitations of the California Psychological Inventory.

Though the California Psychological Inventory (CPI) was acceptable to the advising astrologers, it was complex and limited for calibration. The CPI was more a measure of current outward behaviour [15] than the inner motivations that astrologers identify in a chart. After all, why go to an astrologer if a chart is no more than an analysis based on your own responses? Another limitation was that the "*CPI contains scales that discriminate between the sexes*", Carlson had to withhold the gender of the subjects from the astrologers to avoid clues to identity.

Did the California Psychological Inventory [CPI] fail?

The complexity of the CPI was confirmed when the subjects were unable to identify their own CPI (i.e. the analysis of a personality questionnaire filled in by themselves) any better than chance. Though this result forced Carlson to reject the principle of subject's self-selection [Test #1], the astrologers were better equipped, more informed and more motivated to read the complex CPI tables than the subjects. [12] So the CPI was not an ideal measure, but it proved to have worked to a limited extent in the right hands.

The background to the experiment and many of these criticisms has been well documented by Professor Joseph Vidmar and published in Correlation Volume 26(1) (Vidmar 2009). His concerns about the CPI are reinforced by the CPI test re-test data and the homogeneity of the group gave the astrologers an additional challenge. I do not share Vidmar's view that the experiment failed largely due to the many design faults, procedural errors and inadequate reporting of data – all compounded by the inherent bias. Vidmar believes that Carlson is a unique case of organized scientific misconduct by a group

claiming to "speak for the body politic of science" and even involving the science journal *Nature*. I agree that there were flaws: subject selection, execution faults and the limitations of the CPI. However, these flaws were not fatal, but procedures and techniques that could be adjusted in future experiments of this type. Essentially, the design of the experiment had merit but with scope for improvement.

So though parts of the experiment had to be rejected, the two tests #3 and #4 that were still valid presented the astrologers with the huge challenge of matching CPIs to birth charts.

Carlson's Conclusions

Carlson's conclusions were based on the results of what he considered the two viable tests: Test #3 where the astrologers ranked CPIs against birth charts and Test #4 where the astrologers rated the matches with birth charts out of 10 points. His closing summary stated "*how poorly the astrologers performed when compared to their predicted rate*", "*the astrologers were unable to choose the correct CPI as their first or second choices at a significant level*." and he found "*no convincing evidence that the astrologers tended to rate the correct CPIs higher than incorrect CPIs*." On this basis he concluded "*We are now in a position to argue a surprisingly strong case against natal astrology as practised by reputable astrologers.*" (Carlson 1985:425)

"*The conclusion does not follow from the data*" – Professor Hans Eysenck (Eysenck 1986

In the aftermath of the experiment, a number of authoritative sources including Professor Hans Eysenck of London University (1986) [16] argued that Carlson's conclusion was faulty. By late 2009, new evidence raised the prospect that the *opposite* of Carlson's conclusion may be more appropriate.

Ertel's re-analysis of the data shows a clear trend favouring astrology.

A detailed analysis of the data by psychologist and statistical expert, Professor Suitbert Ertel of the University of Göettingen, Germany in 2009 (Ertel 2009) revealed that the results of the valid tests (#3 and #4) favour astrology to a statistically significant level – in spite of the many disadvantages that the astrologers faced.

Ranking Test #3: Astrologers ranked CPI/Chart matches to a marginally significant level

Carlson's method of analysing 1st and 3rd rank only and separately, missed the overall pattern of an experiment where the astrologers were asked to pick 1st and 2nd choice. "*The astrologers were then asked to select the two CPIs (first* and *second choice, no ties allowed).*" (Carlson 1985:420) Carlson's reductionist approach goes against statistical convention involving choice.[17] His confusing and unconventional use of statistics misses the total effects where the correct horoscope was the most frequent second choice and the last choice was the least frequent. Due to this oversight, his technique overlooked the inherent problem when selecting random CPIs from a subject group where there is little diversity. At least one of the two random CPIs will be similar enough to the authentic one to give an unfair false match with a birth chart. Whenever choices could not be separated, Carlson stipulated that no ties were allowed and so the astrologer (and the subjects) would

often have had to make an arbitrary choice. Carlson's ranking technique would have produced more 'meaningful' results if the astrologers (and the subjects) had been asked to compare dissimilar CPIs. In any event, independent experts should have removed similar CPIs so that every decision could be based on astrology alone.

"I was given some of these charts (CPI profiles) to match myself, and noticed immediately that the three profiles were often quite similar" – Teresa Hamilton (an astrologer who was willing to co-operate but later resigned (Hamilton January 1986)

In test #1 which Carlson rejected as unfair, 73.5% of the 83 subjects identified their correct horoscope as their first or second choice [Chance=66.6%]. In test #3, the astrologers showed slightly superior accuracy in selecting the correct horoscope as first or second choice 74.8% [Chance=66.6%] of 115 [10] subjects. In analysing the total effect using a standard formula for a three-choice format, Ertel showed that the astrologers matched CPIs to natal charts to a *marginally statistically significant* P-value of .054 (Effect Size=.15) [11]. However, more significant results emerged when astrologers rated the charts in test #4. (Ertel 2009)

"Thus. the astrologers chose the correct CPI Profile, either as first or second choice, more frequently than expected by chance, at a marginally significant level." (Ertel 2009, p129)

Table showing Astrologer's Ratings [1 to 10] of matches between authentic & unauthentic CPI profiles with authentic birth charts

Rating 1 to 10	Authentic CPIs	All CPIs	Authentic CPIs as a % of all CPIs	Unauthentic CPIs	Unauthentic CPIs as % of all unauthentic CPIs	Authentic CPIs as % of all authentic CPIs	Authentic CPI % less unauthentic
1	3	18	16.67%	15	7.21%	3.00%	-4.21%
2	10	36	27.78%	26	12.50%	10.00%	-2.50%
3	9	39	23.08%	30	14.42%	9.00%	-5.42%
4	7	27	25.93%	20	9.62%	7.00%	-2.62%
5	16	42	38.10%	26	12.50%	16.00%	3.50%
6	17	38	44.74%	21	10.10%	17.00%	6.90%
7*	13	38	34.21%	24	12.02%	13.00%	0.98%
8*	20	51	39.22%	31	14.90%	20.00%	5.10%
9	4	15	26.67%	11	5.29%	4.00%	-1.29%
10	1	4	25.00%	3	1.44%	1.00%	-0.44%
Total	100	308		208	100.00%	100.00%	0.00%

* The frequencies include 1 additional unauthentic match at rating #7 and 1 additional authentic match at #8 not included in Ertel's published data. Ertel has confirmed these numbers and that this new data does not affect his results.

Rating test (#4): Astrologers rated CPI/Chart matches to a statistically significant level

Test #4 required the astrologers to rate 308 CPI profiles without knowing which ones were the 100 authentic or the 208 unauthentic matches with each birth chart. With this test, each astrologers' confidence in their results could be recorded *much more* precisely and without the ambivalence of having to rank similar profiles arbitrarily as no ties were allowed.

"The correlation is significant. This result gives reason to take into account the probability that the astrologers were able, to some extent, to successfully match birth charts with CPI profiles." Ertel referring to the astrologer's 10-point rating task. (Ertel 2009 p 131)

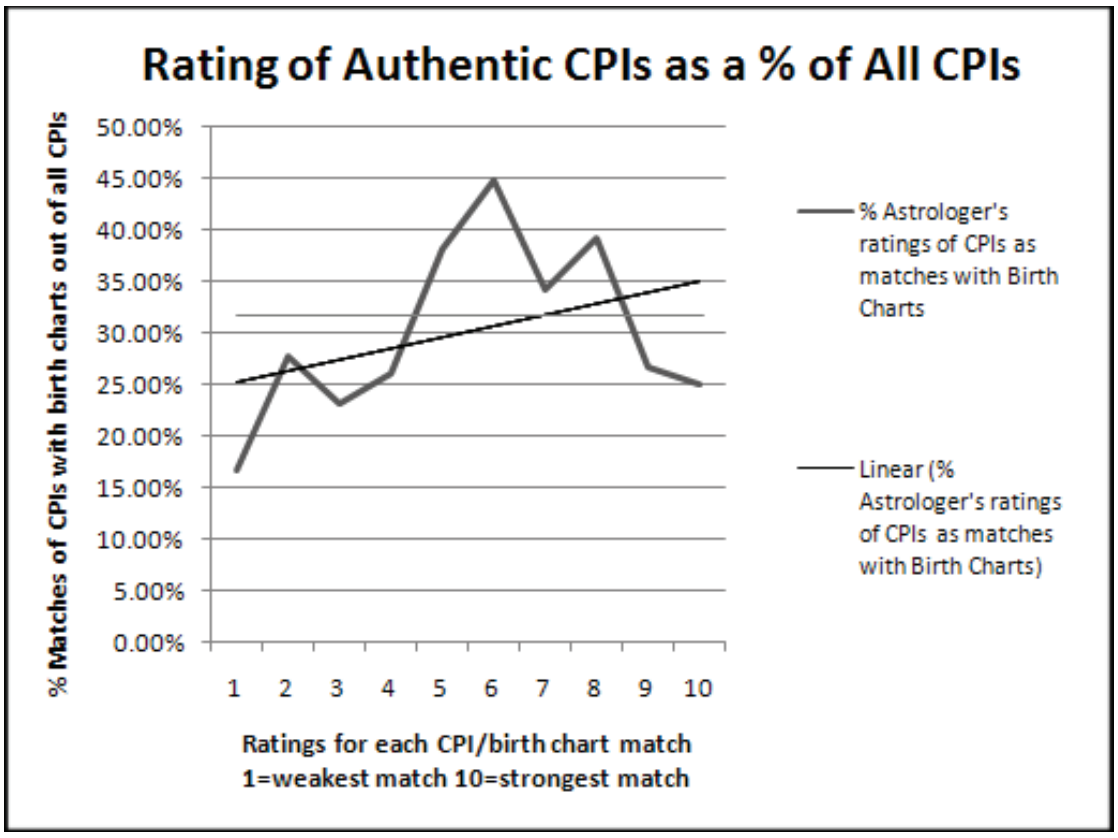
Carlson creates sampling errors by misusing Test #3 data to analyse Test #4

It is hard to imagine how any statistician could justify Carlson's decision to separate this discrete new data [Test #4] into three small samples according to ranking choice from test #3. The two experiments are separate and not comparable; being unrelated in method and sample size: Test #3 114-116 subjects [10] and Test #4 100 subjects. The result was that the problems of the ranking experiment were unnecessarily compounded and the sample sizes much reduced. In his paper in *Nature*, Carlson put up three separate histograms - one for each rank. Although the graphs displaying second and third ranks showed a slope favouring astrology, the first rank graph showed a downward slope, which Carlson accounted for as a chance result rather than a classic sampling error [18]. Paul Kurtz (founder of CSICOP) and his colleagues used this 'sampling error' technique to break down sample sizes in a failed attempt to refute Gauquelin's evidence showing the significant placement of Mars in the charts of top athletes. (Zelen, Kurtz & Abell 1977) (Ertel & Irving 1996) (Rawlins 1981)

Another problem with Carlson's graphs is that there is no allowance for the varying frequency at each rank (except to show a wider range of standard deviation). Therefore the presentation of the data gives undue weight to low frequency results. For example, Carlson's unnecessarily small sample units meant that his slope angle could swing significantly by a single subject intentionally sabotaging the test with hoax answers in his or her CPI encouraging a false match.

On the first graph that follows, the upward trend clearly shows that the astrologers tended to give a higher rating to correct CPIs and a lower rating to falsely matching CPIs in this double-blind test. However, this simple graph is not a truly accurate representation of the results. The reason is that very few birth charts (>20 in each case) were rated at either end of the scale - Levels: 10 (4 CPIs), 9 (15 CPIs) and 1 (18 CPIs). So to take account of the varying frequency, I calculated the percentage of unauthentic CPIs at each rating level of the total unauthentic CPIs (208) and the percentage of authentic CPIs of the total of authentic CPIs (100). The objective was to weight the astrologer's success rate with the frequency. For example, at the maximum rate of 10/10, the astrologers had a 25% success rate (Chance=100/308=32%) but then only rated 4 CPIs. However, they had 39% success rate when rating at 8/10 in evaluating 50 CPIs (12.5 times as many CPIs as the top rate). I then compared the difference between these two percentages to get the weighted difference. The resultant graph (below Figure 2) gives a more realistic picture of how the astrologers performed. The pattern on both graphs should make it quite apparent that *the astrologers (despite disadvantages) were able to discriminate between charts in a*

statistically significant way. This is confirmed by the upward sloping trend line. Ertel calculates that the astrologer's 10-point rating was significant with a P-value of .037 [11] with an effect size of .10. [19]



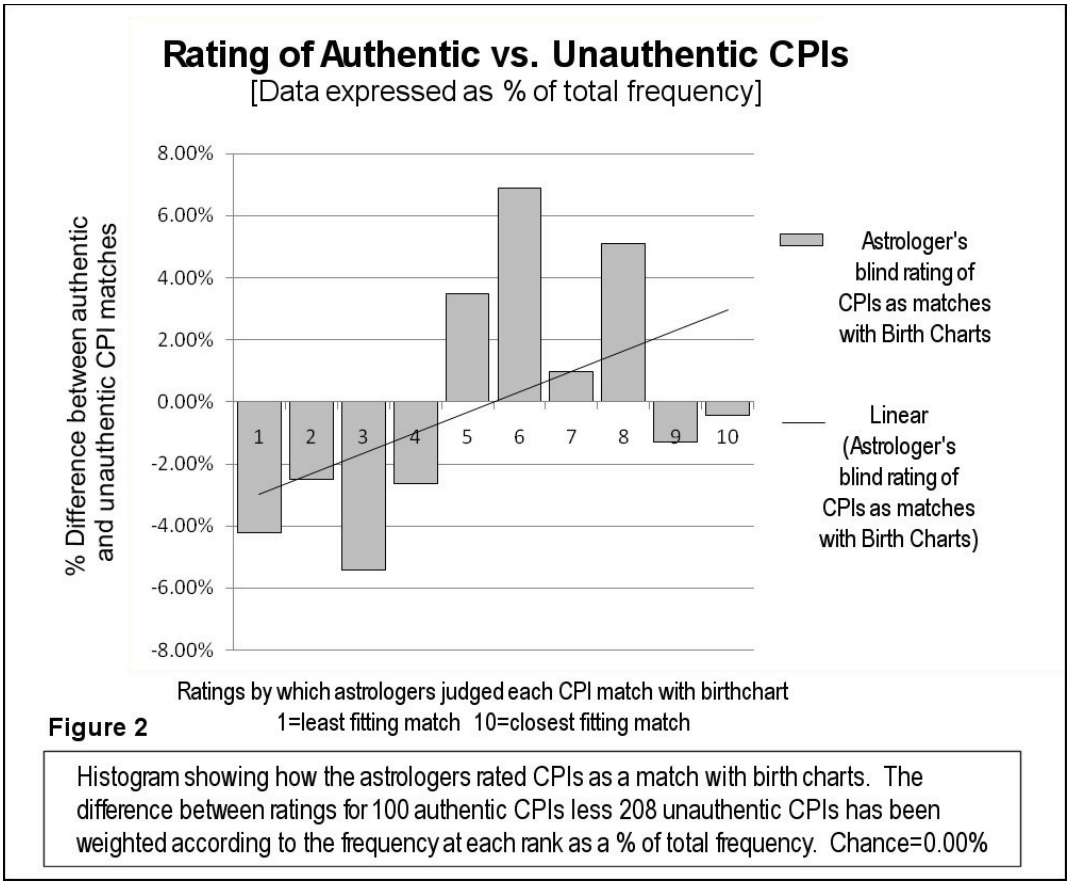


Figure 2

Histogram showing how the astrologers rated CPIs as a match with birth charts. The difference between ratings for 100 authentic CPIs less 208 unauthentic CPIs has been weighted according to the frequency at each rank as a % of total frequency. Chance=0.00%

So in a blind test under unfavourable conditions, the astrologers successfully rated the incorrect matches lower (on the left in Figure 2) and the correct matches between horoscopes and psychological profiles higher (on the right).

How have Sceptics reacted to this Dramatic Reversal?

The Carlson test has yielded powerful evidence for astrology. It has advantages over Gauquelin's extensive studies, which focussed on single chart factors whereas astrologers identify themes from many factors in a chart. The results are not prone to standard criticism of astrological practice: Forer (and Barnum) effects, confirmation bias, cold reading techniques, cherry picking, data artifacts or flattery.

Inevitably, the reaction from sceptics to these new findings that Ken McRitchie (McRitchie 2010) and others have presented on forums has ranged from denial to shock. One sceptic who runs a popular blog dedicated pages to 'authoritatively' debunking Ertel's analysis *before* he had read Ertel's paper. Unlike most astrology experiments, it would be hypocritical for a sceptic to try to debunk the test structure or the procedures. Carlson's experiment was designed to be robust by sceptics using their rules, allegedly 'rigorously peer reviewed' [20] and published in a most prestigious journal *Nature*. It has been proudly upheld as a model experiment and relentlessly defended for almost 25 years as the best evidence that astrology is no more than chance.

Dr Geoffrey Dean's Initial Criticism

In 1986, when Professor Hans Eysenck pointed out Carlson's faulty conclusion, there was no interest from the scientific community or from the Press. In 1989, Carlson claimed, "*I have not yet received a serious scientific challenge to the paper*". Professor Joseph Vidmar (Vidmar 2008) poses that someone like Geoffrey Dean, who advised Carlson on the experiment, must have informed him that the most famous living psychologist at the time [16] had written articles criticising his analysis and conclusion. Dean has remained the most ardent and vocal supporter of Carlson's experiment.

Dr Geoffrey Dean, ex-astrologer, now a CSI (formerly CSICOP) fellow and dogged astro-sceptic is based in Perth, Australia. However he happened to visiting California at the time of the experiment and gave Carlson advice and criticism. I put my case to Dean and he was most helpful and raised a number of objections. His points helped to firm up my opinion of the immense value of the Carlson test and that the reappraisal really is an important 'recent advance in natal astrology'.

Dean commented that the experiment is dated in the sense that later tests with inventories other than the CPI (presumably McGrew and McFall 1990) [21] were more understandable to the subjects. Dean goes on to make a good point that Carlson may have been hasty in dismissing the results from Test #1 since it is easier to identify with an astrological analysis typed in plain English than to interpret the unfamiliar graphical tables on a CPI. So, we can only speculate what would have come from Test #2, as it was unlikely to have suffered from the other limitations in Test #1.

Dean also considers drawing attention to Carlson as a waste of time when there is a "*whole body of research, which currently includes 55 tests of the ability of astrologers to*

match charts to their owners". None of these tests are comparable to Carlson. Of those that I have reviewed including those cited by Dean as the best examples, all appear to have fatal flaws and should to be dropped. Data where the flaw is due to the small sample size like Vernon Clark's (Clark 1961), Nanninga (1996) and Marbell (Marbell 1986) could be combined into a Meta-analysis provided there are no other flaws or artefacts.

Lastly, I understand Dean's point that the results are of little practical value to working astrologers. Even so, I believe that investigating the problems with the experiment and demonstrating that it has yielded evidence will re-open the doors that it had previously closed to future research and acceptance in the academic world.

These are my impressions of several valuable points made in correspondence with Dean. They should not be taken as an exact representation or the limits of his views on the experiment. However, Dean's (and other sceptic's) continued reluctance to accept that Carlson's analysis and conclusions are flawed in the face of what is now overwhelming evidence, suggests that the smaller and less well-designed astrology tests in which he and others claim that astrology has 'failed' have yet to be filtered out by critical analysis.

Was the sample size too small?

Suitbert Ertel, states that the number of recruited subjects and astrologers was insufficient. Ertel's standards are high. He has verified evidence that supports a correlation between planetary positions at birth and professional eminence in studies involving non-participating sample sizes running into thousands [22]. In my view, it is unrealistic to expect a larger sample of participants involving the cooperation of both astrologers and sceptics: Carlson's experiment is the best of only three studies of this type where the sample size is 100 or greater. [23]

Hans Eysenck in his lecture on "*Methodological Errors by Critics of Astrological Claims*" said the study was "*done on a very large scale with many appropriate controls and safeguards.*" (Eysenck 1983) It was large enough for Carlson's decisive conclusion to pass the peer review process and to be published in *Nature*.[20] In addition, in 25 years since publication, an estimated five million people have read or heard the conclusions of this paper.[2] Yet no one appears to have objected to the participant size when the conclusion rejected astrology.

The importance of the number of participants really depends on what is being claimed. To validate the science of astrology would require much more evidence besides what has been accumulated so far. To validate the practice of a few astrologers as scientific, this experiment would need to be replicated with a larger number of subjects but refined by using only the astrologers whose results in a blind selection were above average. Alternative explanations for the results like ESP would have to be ruled out by some means. This is not proof of the entire field of astrology or validation of all astrologers. The results however are sufficient to be considered strong evidence for the successful performance by a few 'reputable' astrologers.

Did Ertel 'mine' the data knowing the outcome? Is this Re-Analysis of the data post hoc?

In 2009, Carlson asserted "*A scientist never figures out how to analyze his data after the fact with all the data in view. This leads to selection biases that can skew the results to favor the experimenter's hopes. Apparently Dr. Ertel has done just that*." (Skeptico 2009)

Carlson's remark appears somewhat hypocritical given that he or one his associates seems to have done exactly what he accuses Ertel of doing. Carlson set out his criteria. However it took nearly four years between the completion of data collection and publication [24]. During that time the data appears to have been re-analysed "after the fact" and re-dressed in a way that favoured Carlson's apparent bias. Ertel has merely done what Carlson originally set out to do and should have done: that is to use standard statistical techniques for the three-way choice #3 and for the rating test #4.

Ertel states that the results are regarded as 'insufficient to deem astrology as empirically verified'. Does this mean that the experiment proves nothing?

Ertel was correct that the results do not 'prove astrology'. This study was not a test of the field of astrology as has been done with the work of Gauquelin and others. It was a test of the practice of astrology by a group of top astrologers who successfully performed to a level that cannot be dismissed as chance.

Surely any evidence can be ignored as the experiment 'does not reject the null hypothesis' (that the astrologers performed no greater than chance)?

In this experiment the null hypothesis is that the astrologer's results are no more than chance. This argument is complicated and technical. The contention is that due mainly to the inconsistent terms and standards of proof that Carlson set for his experiment, the astrologers were theoretically unable to *prove* their results were greater than chance *even though their performance exceeded chance*. But rather than conclude that the experiment is flawed, one imaginative sceptic has tried to claim that this was yet another experiment that merely failed to disprove astrology!

Carlson had framed three hypotheses. The results of the first one of two hypotheses for test #3 did not contradict the null hypothesis. According to Carlson, the "astrologers predicted a correct choice half the time or more" and yet on test #3 the astrologer's first choice match (35%) was not significant. However, by selecting the correct choice 74.8% of the time within the guidelines requested by Carlson below, the astrologers were able to reject the second hypothesis, which was as follows:

"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."* (Carlson 1985:425)

Thirdly, the astrologers' success in predicting their ability to rate the correct CPIs higher than the incorrect CPIs to a statistically significant level disproved Carlson's third 'sceptical' [25] hypothesis.

Besides the evidence this 'null hypothesis argument' has many other faults:

1) First Hypothesis
   i. In his paper, Carlson claimed that his first hypothesis was based on decisive individual selection of single matches. However, this was not Carlson's original instruction to the astrologers for the 3-way match. In the unpublished information sheet to the astrologers, Carlson wrote: *"...Select which profile you feel is (1) most likely, (2) second most likely to be the one."* (Vidmar 2008)

   ii. Carlson claimed he was testing the '*fundamental thesis of natal astrology'*. (Carlson 1985:419) This cannot be measured by a 'confidence judgement' of one or two astrologers on the ability of other astrologers to match unseen birth charts to unknown psychological test results presented in an untried graphical format. No astrologer or psychologist could know what to expect to happen in this unprecedented and unfamiliar test. It would be fairer to describe it as an uninformed guess rather than the first alternative hypothesis and certainly not one that could justify Carlson's unequivocal blanket conclusion.

2) In Carlson's second hypothesis, the astrologers never had an alternative hypothesis to prove.

3) The flaws with the first and second hypothesis are highlighted by the superiority of the third hypothesis. In the third hypothesis, confidence levels were measured with more precision and relevance when the astrologers successfully rated their ability to match CPIs (out of 10) to each natal chart. [Test #4]. The more scope the astrologers were given to predict their judgement, the greater the accuracy with which their decisions could be measured and assessed. As this precision increased, the statistical significance of the results increased. Enabling astrologers to rate each assessment knowing the data (the birth chart and CPI) was an important lesson from the Carlson test and should be considered with all future tests of this type.

Figure 3 (How astrologers performed on 3 hypotheses) shows that as the astrologers were given more scope to express their assessment of the match between each horoscope with the CPIs, their performance became increasingly accurate. The diamond on the left, the result of 1st choice in Test #3 where there were only 3 options, is not significant (little

better than chance). The next result is 1st and 2nd choice in Test #3 where there were 6 options (p=.054). The top and most significant result is Test #4 where the astrologers had one thousand permutations in their choice of 10 rating points for 3 charts. (Of course, more permutations can also make decisions more confusing which can generate random results, but in this case the design empowered the decision making process.)

So the 'null hypothesis' argument appears to be an attempt to cover up facts by exploiting a loophole. Science does not advance through bureaucratic obstruction. It is misleading to cite only the flawed parts of the experiment and even deceptive to exploit an experimenter's faulty and contradictory protocols as a smoke screen to hide evidence.

Conclusion:
Scientific Evidence of Successful Judgements by Astrologers but not Proof of Astrology

Carlson rejected the results from Tests #1 and #2 with some justification. This left only Tests #3 and #4 to verify whether astrology was any more than chance or not.

Without being aware of it at the start of the experiment, the *astrologers faced many disadvantages* in their attempt to identify individual charts. Parts of the tests required arbitrary guesswork, which is not part of standard astrological practice. The demographic profile of the subject group was too uniform to make individual identification easy. The subjects were not typical of astrologers' clients being less mature, more male, less motivated towards self-enquiry and mostly disinterested in applying themselves to what was a challenging test. The astrologers had to work with a Psychological Test (CPI) that was complex, subjective and produced any limited ephemeral data that were unlike the more powerful information that can be obtained from a birth chart.

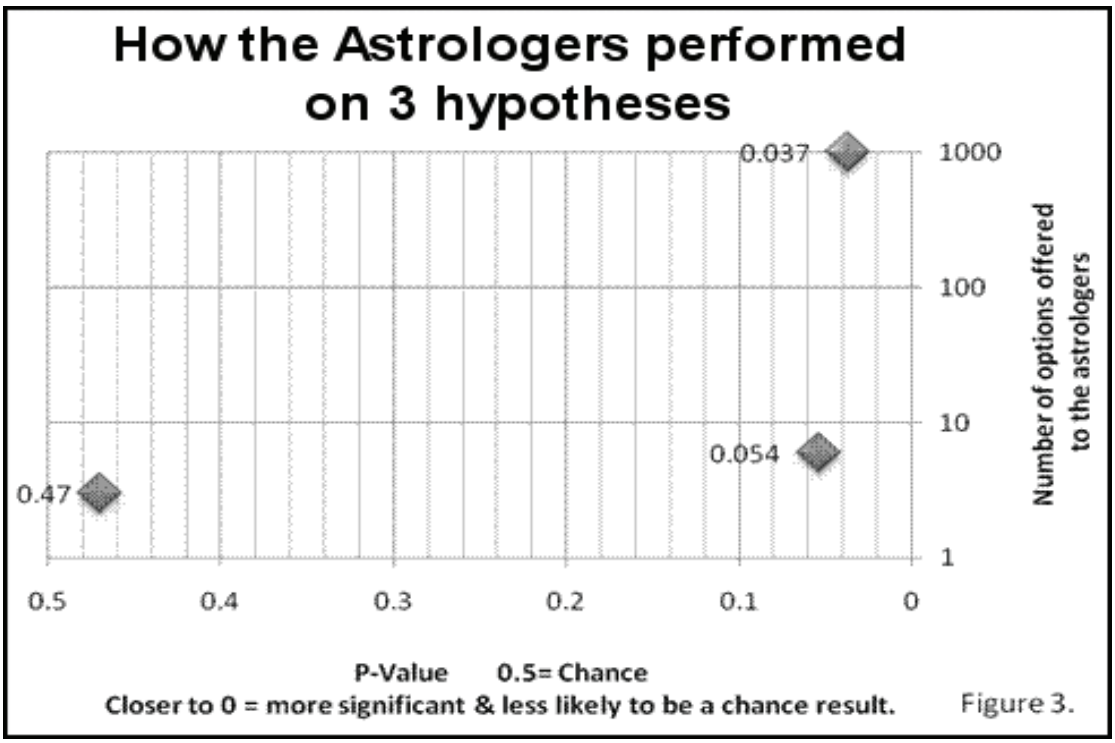


Figure 3.

However, while these handicaps impaired the astrologers' decision-making in test #3, the astrologers were given ways around them in test #4 (by rating poor matches lower, for example). So, by applying standard statistical methods to Carlson's original criteria in what was the most realistic test (#4), Professor Ertel was able to show that the astrologers were able to rate natal chart matches (out of 10) in a blind test with what appeared to be scientific accuracy. The probability that this result occurred by chance is 37 in 1000 (p=.037) or the unlikely chance of tossing a coin twelve times and getting the same coin face at least ten times. As yet, there are no strong arguments to refute this claim.

Why this significant result took over 20 years to uncover was mainly down to the sampling errors in the preparation and presentation of the data. Sampling errors have also appeared in large tests supporting astrology by Gauquelin - found by Ertel in 1987 (Ertel 1988) and later corrected - and in the Sun sign data analysis (Sachs 1999) compiled by German multi-millionaire mathematician and entrepreneur, Gunter Sachs. Sachs' data needs to be reanalyzed. Sun sign tests are not a test of the practice of astrology and limited by the subject's prior knowledge of their sun sign. While these tests have shown only limited significant correlations without this knowledge, tests of the whole chart such as Clark (1961) and Marbell (1986) have yielded significant results though Nanninga (1996) did not. Unlike Carlson these tests were limited by the small sample sizes and the astrologers had no scope for confidence rating of their judgements.

However, the Carlson Test has now yielded scientific evidence, which needs to be taken into account along with other studies supporting natal astrology including Gauquelin (1994), Ertel & Irving, "*Tenacious Mars Effect*", (1996), Timm & Köberl (1986), Müller & Menzer (1993), Addey (1994), Hill (1996), Douglas (2001), McGillion (2002), Harris (2008), McMahon (2010), financial astrology: Zheng (2001) and Pelc (2010) and natural astrology such as the tides: originally from Ptolemy, *Tetrabiblos* (ca.150), Kepler, *Astronomia Nova* (1609) and Newton, *Principia* (1726) et al, weather: Ding (1982), Cerveny (1997-2010), Varshneya (2010), earthquakes: Tamrazayan (1968), Zhao, Han and Li (2000) and Johnston (2008) and the Sunspot/Planet/Life interrelationship: Brown, Webb & Bennett (1958), Seymour (1997), Wainwright (2004), Hung (2007) and Wilson (2008).

These are empirical studies that fall within the realm of astrology: the study and ongoing search for correlations between celestial phenomena and life, events and physical processes on Earth. Lest anyone questions the range of astrological studies listed, throughout most of the first millennium BCE, Babylonian astrologers recorded Diaries (Menologies) which systematically contained celestial phenomena along with coinciding mundane information such as weather, water levels in the Euphrates river, prices of commodities, political or unusual events such as earthquakes. Their empirical work led to the world's first database. (Campion 2008)

Many of these modern scientists listed above might not consider themselves astrologers and some of these astrologers might not consider themselves scientists. Yet, these researchers strive to follow the scientific method in their quest to identify cosmic

connections, patterns and cycles. Their conclusions are falsifiable and their data open to evaluation.

While the practice of astrology is mostly outside the realm of science, the only justification to dismiss the entire field of astrology as rubbish (from a scientific point of view) or as a pseudoscience is ignorance. The oldest science in the world still appears to have hard science at the core and yet few of its practitioners can and do claim to be scientists. The results of the Carlson Double-Blind Astrology Test itself does not prove the validity of astrology, but add to the growing body of evidence that the practice of professional astrologers and astrology should once again be taken seriously as a field worthy of study.

"Let the dataset change your mindset" – Professor Hans Rosling, Swedish Public Health Scholar (Rosling 2009)

Remaining Thoughts and Questions

How did a scientist with such a strong academic background reach such an unfounded conclusion?

Since his experiment, Carlson has acquired a Ph.D. in nuclear physics and is founder of the Society of Amateur Scientists. After the alleged skulduggery (Rawlins 1981) and dogmatism (Lippiard 1995) involving CSICOP and other sceptical organisations attempting to debunk Gauquelin's evidence (in which Carlson was not involved). It might reasonably be wondered whether inconvenient data was conveniently ignored. Perhaps a clever statistician from CSICOP found a way to 'dumb down' the presentation of the data and guided Carlson's youthful inexperience towards an approved conclusion? Unless proved otherwise, I take an optimistic view of human nature.

Carlson's untenable conclusion may be more a reflection of his mindset than any sleight of hand. He appears to operate by simplifying and reducing data to numbers and statistics. His training as a physicist would be in quantitative rather than qualitative analysis. By involving human behaviour, astrology requires both approaches. It's as if Carlson was looking for black and white answers whereas human nature can only be understood in colour. This reductionist mindset – often found in sceptics – is the inability to see patterns or consider that the whole can be greater than the sum of the parts. While this linear thinking approach may be effective in the science of inanimate physical data, it tends to miss the point in the social sciences or the humanities or what I would term the 'animate sciences'.

Why did *Nature* uncritically accept this study for publication?

The key may lie in the fact that the editor: John Maddox, was a CSICOP fellow and this may have influenced his judgement of the research making it difficult for him to objectively critique and evaluate its methods and conclusions. Maddox later described the experiment and how the astrologers drew up "*Lunar Charts"* (sic) for "… *a perfectly convincing and lasting demonstration*". Elsewhere he said "*It is a plain fact that astrology is a pack of lies in the literal sense*." (Maddox 1994) However, when it came to scientific objectivity, Maddox appears to have given priority to publishing an experiment that he

hoped would confirm his deeply held beliefs over critical thinking. Even so, astrologers can thank Maddox. Had Carlson submitted the correct analysis of the data showing support for astrology, it is inconceivable that Maddox would have considered it for publication in *Nature*.

What was Carlson's conclusion about the role of psychology?

If any field fell short of Carlson's expectations in this exercise, it was psychology (and not astrology) with the perceived 'failure' of the subjects to identify their own CPI. Carlson however, came up with four excuses (Carlson 1985:425) to exonerate the CPI. Although he could have excused the astrologers in the same way, Carlson was less merciful towards astrology. And as a result of his bias, astrology – instead of psychology - unjustly took the rap from the Press.

Could the results improve astrological research?

Since the rating of psychological profiles by astrologers proved viable under difficult conditions, astrological schools might consider developing their own 'Astro Psychological Profile' where mature, open-minded subjects from diverse areas answer a questionnaire designed to measure features that astrologers might typically find in a birth chart. Students could then be tested by their ability to rate the profile as a match with a birth chart. This might also be a way to refine techniques, discover more about the strengths and limitations of astrology and become a good teaching tool.

What is the status of this experiment?

It's ironic that Wikipedia and many sceptical websites undermine their often-questionable credibility by citing the Carlson experiment as one of the strongest arguments against astrology, when the results now favour astrology. As news of this reversal of the conclusion spreads, sceptics will be put in an invidious position where they must decide whether to continue to support the test and risk validating astrology or to claim that the test is fatally flawed. In the absence of conclusive scientific experiments refuting modern astrology [26], a chasm has now opened. Astrologers can now legitimately cite Carlson's study published in the most prestigious science journal *Nature* as further evidence supporting the thesis that while the practice of astrology is predominantly an art, the field of natal astrology increasingly appears to have some scientific basis.

Robert Currey, BSc.[Hons], DF Astrol.S., Cert.Int.ACG.

Acknowledgements

I wish to thank Dr. Pat Harris, Ken McRitchie, Dr. Geoffrey Dean, Professor Suitbert Ertel, Professor Joseph Vidmar, Karen Chiarello, Deborah Houlding, and Jill Davies for their help and advice.

Endnotes

[1] Ofcom ITC programming code which binds the BBC, (according to their own publication) reads: Horoscopes, palmistry and similar 'psychic' practices are only acceptable where they are presented as entertainment or are the subject of legitimate investigation. (Ofcom

2010) Cox's comments were not part of a legitimate investigation and they were not intended for entertainment

[2] "Carlson's claims that 'astrology failed' and 'there is no scientific evidence for astrology' have been heard or read by over an estimated 5 million people." (Vidmar 2008)

[3] CSICOP stands for Committee for Scientific Investigation of Claims of the Paranormal is now renamed CSI (Committee for Skeptical Inquiry). Brian Cox is on an editorial Advisory Board for the Skeptic newsletter issued by a group associated with CSICOP. (Skeptic 2010)

[4] "The years 1981-1982 were years of crisis for CSICOP and many of its senior members resigned in protest over the Gauquelin affair." (Vidmar 2008) For more evidence of CSICOP involvement and bias: (Currey October 2010) and for an inside story from a co-founder of CSICOP: *sTARBABY* (Rawlins 1981)

[5] Paul Kurtz claimed CSICOP encouraged Carlson to do this project. (Kurtz 2006) Richard Muller (CSICOP fellow) funded and guided the experiment. (Carlson 1985 p.425) and The Editor of the Journal, *Nature*, which published the experiment was John Maddox (CSICOP fellow). (SI July/Aug 2009)

[6] On a listing on a library site on the web, Carlson is listed as being born on 11 March 1960. (Library Thing 2010)

[7] In newspaper interviews, Carlson admits that he worked his way through college as a street psychic and professional magician. (Vidmar 2008)

[8] In his *'Astrology Experiment Detailed Outline'*, Carlson wrote he intended to publish the experiment in the *Skeptical Inquirer*, a "*nationally respected scientific journal*". This unmerited glowing description of the in-house journal for CSICOP, dedicated to debunking the type of experiment he was supposedly hoping would be favourable to astrology, does not fit the profile of a neutral party. Unlike scientific and astrological journals, the Skeptical Inquirer "*... nearly always presents only one side of a controversy in its articles.*" (Hansen 2009)

[9] Though Carlson refers to it as the California Personality Inventory throughout his 'peer-reviewed paper, Professor Harrison Gough the inventor (who was apparently consulted during the test) named it the California Psychological Inventory- also referred to as the CPI.

[10] Carlson's figures for test #3 show a discrepancy with totals of 116 1st choice, 114 2nd and 114 for 3rd choice. I have calculated on the basis of 115.

[11] The P-value is a measure of the reliability of a result. It is an index of the probability of an observed difference in results if the samples compared were random. So a P-value of .5 or 50% suggests the relationship is no more than chance. A P-value of .05 indicates that there is a 5% or 1/20 probability that the observed result is a fluke. P-values of .05 or less are usually considered indicative that the results are statistically significant in social sciences and in this experiment a strong likelihood of a correlation. (Thisted 1998)

[12] The subjects were not predisposed to astrology, unlike a typical client that might consult an astrologer. To avoid bias, Carlson eliminated those who strongly disbelieve in astrology and anyone who had had a birth chart prepared before. (Carlson 1985 p.421)
Subjects were less motivated and potentially subject to response bias: According to Ertel, there are two possible reasons why the subjects were not as good as the astrologers in matching their charts or in completing their CPIs: (Ertel 2009)

a. While the astrologers put in great effort as '*their world views were put on stake*', the subjects were not motivated. This is evident by the fact that "*half their data (ratings) were so poor that they could not even be analyzed.*"
b. The subjects might have been inclined to avoid the correct chart interpretation (or their CPI) to deny character traits they might consider unattractive.
Astrologers were more informed: According to Vidmar, "*The astrologer subjects were given the entire 28-page 'Interpreter's Syllabus' to assist with CPI interpretation and the student (sic) subjects were given only a 2-page 'synopsis' or 'summary' of what the individual scales meant" (Vidmar 2008)*

[13] Carlson's paper in *Nature* claims 70% were college students and about half were graduates. No other demographics were given though in 1986 Carlson in response to questions claimed that the mean age of the subjects was 28. (Carlson 1986)

[14] According to CPP - publishers of the CPI "Test-retest correlations for high school students over a one-year interval range from .52 to .73 with a median of .66. Test-retest correlations for adults over a 10-year interval range from .49 to .85 with a median of .77."
"The reliability of the CPI has been assessed as to its internal consistency, as judged by ... test-retest (medians: 1 year=. 68, 5 year=.56, 25 year .58)." (Info Refuge 2010) To amplify this point about the evolution of the personality, test-retest results of the CPI after 25 years shows a median of .58. Or to put it another way after 25 years, half the subjects had an altered perception of their self image by a factor of 42% according to the CPI measuring system. Since self-knowledge tends to improve with age, we can only assume that this adjustment is to greater reflect their true selves.

[15] Harrison Gough, author of the CPI states: "*Each scale is designed to forecast what a person will say or do under defined conditions, and to identify individuals who will be described in characteristic ways by others who know them well or who observe their behaviour in particular contexts. The scales are grouped for convenience into four broad categories, bringing together those having related implications. The underlying logic here is interpretational, not factorial, i.e., these four categories do not necessarily constitute psychometric entities."* (Gough 1987)

[16] Eysenck, [b.4 March 1916 – d.1997] was Professor of Psychology at the Institute of Psychiatry, authored about 80 books and at the time of his death was the living psychologist most frequently cited in science journals. (Haggbloom 2002) Eysenck's expertise was in measurement of personality. He was one of very few psychologists who researched the field of astrology.

[17] Ertel (2009) p.128 "*He (Carlson) should have used pair comparison, the fairest existing test format on which Thurstone's "law of comparative judgment" had been based."* (Thurstone, 1927)".

[18] In statistics, sampling error or estimation error is the error caused by observing a sample instead of the whole population.

[19] Effect size is a standardized measure of the strength of the relationship between the two variables. This descriptive statistic (magnitude) can complement the p-value (probability).

[20] Since Carlson's experiment was published in the Commentary Section in *Nature*, there was a question mark over the peer review. Carlson responds to this "It survived a rigorous peer review that included a famous psychologist whom I will reveal in a later publication. (Carlson 2009)

[21] (McGrew & McFall 1990) This experiment appears to be a genuine attempt to overcome the Carlson limitations of the CPI by providing astrologers from a local group in Indiana, an abundance of information. Unfortunately, this included photographs that meant that the age range 30-31 had to be very narrow (to avoid clues from the photos). Though the subjects appeared to have diverse lives, they were all born around the Saturn Neptune conjunction. So there was a higher level of homogeneity than Carlson's subjects and a similar problem with lack of self-awareness/experience due to age. However, the unrealistic challenge was that the astrologers had to match each birth chart to not 3 CPIs but, 23 psychological profiles! In addition, the astrologers were asked for first choice only, so there was no opportunity to rate and reject poor matches, which dramatically increased the power of the Carlson experiment. Though there are no data tables to analyse, the experimenters assure that the astrologers were unable to match significantly and were surprised at their lack of agreement. This result was inevitable. McGrew & McFall designed an impossible test that favoured random results in a small homogenous sample.

Their paper was critical of Carlson's use of the CPI and his conclusions. Perhaps it is not surprising that it was rejected for publication in *Nature* as it was Carlson who did the peer review. Given *Nature*'s history of only publishing flawed but negative tests on astrology, rejection of an astrology study by *Nature* may not be a reflection on the quality of the experiment.

[22] Michel Gauquelin claimed that Mars was located in specific sectors in the birth charts of sports stars from a database of 4,384. Ertel was able to verify this by comparing it with data compiled by 3 sceptical groups (N=1664) by showing a trend of increasing significance according to the sports champion's eminence measured by independent citation. (Ertel, S. & Irving, K., 1996),

[23] Dean states that "*so far only three tests of this kind have exceeded N=100, one of them Carlson's."* (Dean 2010)

[24] Data collection was completed on 15 Dec 1981. The paper was submitted to *Nature* on 11 March 1983. *Nature* published the paper on October 14 1985. Why was there a gap of 3 years and 10 months? (Vidmar 2009)

[25] What Carlson termed the scientific hypothesis is a misnomer since you cannot disprove a scientific hypothesis. So to avoid confusion, I have termed it the sceptical hypothesis.

[26] For example magic tricks, anecdotal evidence, Sun Sign only data or tests with sample sizes that are so small that they are likely to produce random results would not be considered valid tests.

References

Addey. John (1994), *Harmonics in Astrology*. Eyebright Books. Somerset, UK. 1994.

Brown, Webb & Bennett, (1958), *Comparisons of Some Fluctuations in Cosmic Radiation and in Organismic Activity During 1954, 1955 and 1956,* American Journal of Physiology*,* Dept. of Biological Sciences, Northwestern University, IL; Dept. of Physiology, Goucher College, MD; Dept. of Biology, Sweet Briar College, VA; Marine Biological Laboratory, MA, http://ajplegacy.physiology.org/content/195/1/237.abstract

Campion, Nicholas (2008), *A History of Western Astrology*, Vol.1 p.75, Continuum Books, London.

Carlson, Shawn (2009), comments on web on article "*Astrology Still Fails*" by Skeptico. August 3rd 2009 http://skeptico.blogs.com/skeptico/2009/08/shawn-carlson-astrology-test-nature-suitbert-ertel-reappraisal.html

Carlson Shawn (1985), *A double-blind test of Astrology*, *Nature* Vol.318, pp.418-425 [5 December 1985]

Carlson,Shawn (1986) *Summary of Carlson's Double-Blind Test of Astrology* by Dean. or possibly Smit, R .Article date and author are not specified though the context suggests Dean. http://www.astrology-and-science.com/d-rese2.htm checked 5 May 2011

Cerveny, R. S. Shaffer, J. A., & R. C. Balling Jr. (1997), *Polar temperature sensitivity to lunar forcing?,* Geophys. Res. Lett., 24(1), 29-32.

Cerveny, R. S., and R. C. Balling Jr. (1999), *Lunar influence on diurnal temperature range*, Geophys. Res. Lett., 26(11), 1605-1607.

Cerveny, R. S., and J. A. Shaffer (2001), *The Moon and El Niño*, Geophys. Res. Lett., 28(1), 25-28.

Cerveny, R. S., B. M. Svoma, and R. S. Vose (2010), *Lunar tidal influence on inland river streamflow across the conterminous United States*, Geophys. Res. Lett. , 37, L22406, doi:10.1029/2010GL045564.

Clark, Vernon (1961), Experimental astrology, Aquarian Agent, 1(9)

Cornelius, Geoffrey, (1994*), The Moment of Astrology*, Appendix 2 (pp.189-193), Penguin Arkana

CPP (2010), *History of the CPI 260® Instrument*, ttps://www.cpp.com/Products/cpi260/cpi260_info.aspx retrieved July 2010.

Currey, Robert (Aug.2010) *Secrets behind a Test of Astrology by illusionist Derren Brown*, http://www.astrologer.com/tests/dbtest.htm

Currey, Robert (July 2010) *Illusionists are for entertainment, not to masquerade or undermine science*. http://www.astrologer.com/tests/randitest.htm

Currey, Robert (October 2010), *Was CSICOP scientific and is CSI merely sceptical?,* www.astrologer.com/tests/csicop.htm

Dean, Geoffrey (15 July 2010), Email to Robert Currey
Dean, Geoffrey (1997) *Meta-analyses of nearly 300 empirical studies* http://www.astrology-and-science.com/d-meta2.htm, last updated 1997 (accessed 4 March 2011).
Dean. G and Kelly, I (2003), *Is Astrology Relevant to Consciousness and Psi?, Journal of Consciousness Studies,* www.imprint.co.uk/pdf/Dean.pdf p.188
Ding & Reiter,(1982) *A relationship between planetary waves and persistent rain- and thunderstorms in China,* Archives for Meterology, Geophysics and Bioclimatology, Ser.B, 31,pp. 221-252, Dept. of Atmospheric Science, Colorado State Uni., CO
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, S. & Irving, K., (1996), *Tenacious Mars Effect*, Urania Trust
Ertel, Suitbert (2004), *Hopeful Findings, Unduly Neglected, on Stars and Human Affairs*, Correlation, 22, pp.4-19 (2004) and Australian Journal of Parapsychology, 2007, Vol.7, Number 1, pp.52-71
Ertel, Suitbert, (2009), *Appraisal of Shawn Carlson's Renowned Astrology Tests*, Journal of Scientific Exploration, Vol.23, #2. pp.125-137
Eysenck, H. J., & Nias, D. K. (1982). *Astrology: Science or superstition?*, London: Temple Smith; New York: St Martin's
Eysenck, Hans (1983), Methodological Errors by Critics of Astrological Claims. Opening lecture at 3rd Institute of Psychiatry Conference, May 1983
Eysenck, Hans, (1986) "*Critique of a 'Double-Blind test of Astrology*"', Astro-Psychological Problems, Vol.4, January 1986
Facebook (2010), http://www.facebook.com/astrology, edited by Robert Currey, Equinox Astrology, [April 2010]
Frazier, Kendrick (2009) "*John Maddox, Long time Nature Editor and CSI Fellow Dies"*, by Skeptical Inquirer Volume 33.4, July / August 2009
Gauquelin, Michel (1988), *Written in the Stars*, Aquarian, Press, Wellingborough, Northamptonshire, UK
Gauquelin, Michel (1988), *Is there a Mars effect?* Journal of Scientific Exploration, Vol.2, 1
Gauquelin, Michel (1994), *Cosmic Influences on Human Behavior: The Planetary Factors in Personality*, Aurora Press. Santa Fe, NM.
Gough, Harrison G. (1987), *California Psychological Inventory Administrator's Guide*. Palo Alto, CA: Consulting Psychologists Press, Inc. & California Psychological Inventory (CPI) Manual, (1996), Consulting Psychologists Press, Inc.
Haggbloom, S.J. (2002). *The 100 most eminent psychologists of the 20th century.* Review of General Psychology, 6, pp.139–152.

Hamilton, Teresa, (April 1986), *Review of Carlson's Experiment*, Astro-Psychological Problems, Quarterly Research Journal,
Hamilton, Teresa, (January 1986), *Critique of the Carlson Study*, Astro-Psychological Problems, Quarterly Research Journal, Vol.4 (1) & Kosmos (1986)
Hansen, George. (1992). *CSICOP and the Skeptics: An Overview* The Journal of the American Society for Psychical Research, Volume 86, No. 1, January 1992 p.24
Harding, Mike "*Astrology in the Year Zero*": http://www.astrozero.co.uk/astroscience/harding.htm#_edn15 *"Prejudice in Astrological Research*" & Correlation Vol 19 (1) 2000
Hill, Judith, (1996), *The Mars Redhead Files*, Stellium Press, http://www.judithhillastrology.com/Stellium-Press-Books/THE-MARS-REDHEAD-FILES?pop=0
Hung, Chin Cheh (2007), *Apparent Relations Between Solar Activity & Solar Tides caused by Planetary Activity*, NASA, Glenn Research Center, Cleveland, Ohio
Info Refuge (2010), *California Personality Inventory*, http://www.inforefuge.com/california-personality-inventory retrieved July 2010
Johnston, Brian T. (2001), Planetary Aspects and Terrestrial Earthquakes, Cura
Kepler, Johannes, *Astronomia Nova* (1609) via Voelkel, James, *The Composition of Kepler's Astronomia Nova.* (2001), Princeton, p.72
Kurtz, Paul (2006) *Science and the Public: Summing Up Thirty Years of the Skeptical Inquirer*, *Skeptical Inquirer,* September 2006 via Skeptical Inquirer Vol.33.4 July/Aug 2009
Library Thing (2010) http://www.librarything.com/commonknowledge/search.php?q=1960-03-11&f=8&exact=1 (accessed 10 Sept 2010)
Lippiard, Jim (1995), *Foreword to The Tenacious Mars Effect, pp.iii-iv, Urania Trust, 17 Dec.1995*
Marbell, Novak, Heal, Fleming & Burton (1986), *Self Selection of Astrologically Derived Personality: An Empirical Test of the Relationship Between Astrology and Psychology*, NCGR Journal, Winter 1986-87, pp.29-44
McGrew & McFall, "*A Scientific Inquiry into the Validity of Astrology*", 1990.
McGillion, Frank (2002), *The Pineal Gland And The Ancient Art Of Iatromathematica*, Journal of Scientific Exploration Vol. 16, No. 1 pp 19-38
McRitchie, Ken (2010), *Reappraisal of the Carlson study finds support for astrology* ~ http://theworldedge.blogspot.com
McRitchie, Ken (2011), *Support for Astrology from the Carlson Double-Blind Experiment,* Astrology News Service, http://astrologynewsservice.com/articles/support-for-astrology-from-the-carlson-double-blind-experiment/

McMahon (2010), McMahon, Ciarleglio, Axley, Strauss, Gamble: *Perinatal photoperiod imprints the circadian clock,* Nature Neuroscience, 14, 25–27 (2010) http://www.nature.com/neuro/journal/v14/n1/pdf/nn.2699.pdf
Maddox, John, (1994), *Defending Science Against Anti-Science*, *Nature*,Volume 368, Issue 6468, pp. 185
Maddox, John, (1995) *An Encyclopedia of Claims, Frauds, and Hoaxes of the Occult and Supernatural*, http://www.randi.org/encyclopedia/astrology.html retrieved Aug 2010 James Randi Foundation, author and date not listed. 1995 estimated. Marked Copyright (C) 1995-2007 James Randi
Muller, Richard (2010) Professor of Physics at University of California at Berkeley. http://muller.lbl.gov (viewed 26 July 2010.)
Müller & Menzer (1993), 1145 *Angehörige deutscher Dynastien* [1145 members of German dynasties]. Astro-Research Data 4. Waldmohr, A.P. Müller-Verlag (via Ertel 2004)
Ofcom ITC programming code which binds the BBC, (according to their own publication) http://www.ofcom.org.uk/static/archive/bsc/pdfs/research/occult.pdf p4 para3.
Nanninga, Rob, (1996),*The Astrotest, A tough match for astrologers*, Correlation, Journal of
Research into Astrology, Winter 1996/97, 15(2), p. 14-20.
Newton, Isaac, (1726) *Philosophiae Naturalis Principia Mathematica,* translated from Latin by Cohen, Bernard, (2000) The Principia: Mathematical Principles of Natural Philosophy, University of California Press.
Nitsche M. E. (2002), *Are the Stabilizing and Destabilizing Influences of the Planetary Gravitational Field on the Structural Formation of Complex Systems Real?- Triggering of Earthquakes*, Lecture at 2002 Annual Conference of the International Association for Mathematical Geology; Berlin, Germany
Pelc, T & Bondar, Dmytro (2010), *Sheer Lunacy staring at the Heavens,* Charting Equity Special, Royal Bank of Scotland, 7 July 2010
Plait, Phil, (2010) *Astrologers Jump on Cox*, http://blogs.discovermagazine.com/badastronomy/2010/04/07/astrologers-jump-on-cox/ April 7th, 2010 12:12 PM
Ptolemy, Claudius, (ca150 CE) *Tetrabiblos*, edited by F. E. Robbins, Harvard University Press, Loeb Classical Library. 1940.
Rawlins, D (1981), *sTARBABY*, Fate, No: 34, http://cura.free.fr/xv/14starbb.html
Rosling, Hans, (June 2009), *Let my Dataset change your Mindset*, Ted Video, http://www.ted.com/talks/hans_rosling_at_state.html
Sachs, Gunter (1999) *The Astrology File: Scientific Proof of the Link Between Star Signs and Human Behaviour*, Orion Books -December 1999.
Seymour, Percy (1997), *The Scientific Basis of Astrology*, W. Foulsham, Slough, U.K.: Quantum, November 1997

Skeptico, (August 03, 2009), *Astrology Still Fails,*
http://skeptico.blogs.com/skeptico/astrology/
Smit, Rudolf, (2010) [undated, viewed on web 25 June 2010] Research Results - Summary of Carlson's Double-Blind Test of Astrology, http://www.astrology-and-science.com/d-rese2.htm
Snow, Ed (2009), *Famous Test Of Astrology Is Seriously Flawed*, EIN Presswire, www.einnews.com/pr-news/36002-famous-test-of-astrology-is-seriously-flawed, 14 July 2009
Tamrazyan, G. (1968), *Principal Regularities in the distribution of Major Earthquakes Relative to Solar and Lunar Tides and other Cosmic Forces*, Icarus, Vol.9,pp.574-592, Inst. Of Geology, Baku, (former USSR)
Thisted, Ronald, (1998) *What is a P-Value?,*
http://www.stat.uchicago.edu/~thisted/Distribute/pvalue.pdf p5.
Thurstone, L (1927), *A Law of Comparative Judgement.* Psychological Review, 34 pp.273-286
Timm & Köberl (1986), Re-analyse einer validatätsuntersuchung, an 178 astrologen, Zeitschrift für Parapsychologie und Grenzgebiete der Psychologie, 28,33-55 (via Ertel 2004)
Varshneya, Vaidya, Vyas PandeyChimote, Damle, Shekh & Karande, (2010), *Forecasting Rainfall for 2010 of Gujarat based on Astro-meteorology*, Anand Agricultural University, Anand – 388 110, Gujarat, India
Vidmar, Joseph (1987) letter from Teresa Hamilton Weed
Vidmar, Joseph, (2009) *Irregularities in the Publication of the Carlson Paper* via email to Robert Currey.
Vidmar. Joseph, (2008) *A Comprehensive Review of the Carlson Astrology Experiments,* Correlation, Volume 26 (1)
Wainwright, Glyn, (2004), *Jupiter's influence*, New Scientist Issue 2439, (20 March 2004)
http://www.newscientist.com/article/mg18124394.100-jupiters-influence.html
Wikipedia, (2010), *Astrology*, subsection Research
http://en.wikipedia.org/wiki/Astrology#Research [viewed 13 September 2010]
Wikipedia, (2011), "*Aquirata (talk contribs), Petersburg (talk contribs), Costmary (talk contribs), Erekint (talk contribs), Apagogeron (talk·contribs), and Gary PH (talk contribs) are all banned from Astrology and related pages per [97]. Moreschi (talk) 16:51, 21 March 2011 (UTC*)"
http://en.wikipedia.org/wiki/Wikipedia:Requests_for_arbitration/Pseudoscience#Final_decision The stated reason for the ban was that these expert astrological editors had only edited astrology and astrology related pages on Wikipedia and were considered 'Single Purpose Accounts'.

Wilson, I., (2008), *Do periodic peaks in planetary tidal forces acting upon the Sun influence the sunspot cycle?* Faculty of Sciences, University of Southern Queensland, QLD

Zelen, Kurtz, Abell (1977), *Is there a Mars Effect?* The Humanist, Nov-Dec 1977, p.36-39

Zheng, Lu, Yuan, Kathy Zhichao and Zhu, Qiaoqiao, (2001), *Are Investors Moonstruck? - Lunar Phases and Stock Returns* (September 5, 2001). Journal of Social Science.

Zhao, Han & Li (2000), *Variation of Lunar-Solar Tidal Force and Earthquakes in Taiwan Island of China*. Earth, Moon, and Planets (Springer Netherlands) 88 (June, 2000) pp.123–129