Ertel's 'Mars Effect': Anatomy of a pseudo-science

Jan Willem Nienhuys

In my article I merely mentioned the null result of CFEPP's study. Suitbert Ertel attacks this study severely, so I will discuss that study in some more detail. However, my silence on some of the innumerable statements in Ertel's contribution is merely because of lack of space, it doesn't mean I silently acknowledge that he is right, as Ertel (1995b) has inferred before. I will conclude my answer with the reasons why I see no point in further discussion with professor Ertel.

If one believes Ertel, the most important part of the design of the CFEPP test was that the selection of the champions had to be performed by people who didn't know about the purpose of the selection, if necessary by hired students. Now professor Kurtz hired students for the US test (Kurtz et al., 1980), and one can read (Ertel and Irving 1996, p.A2-26) about repeated efforts many years afterwards to find out from them personally why the US test might have produced a null result. After the test is over, it's always easy to tell stories about the never-never-land of might-have-been. I propose to stick to facts.

Ertel misrepresents the agreement between CFEP and Gauquelin

The 1982 protocol of the CFEPP test said: 'In order to eliminate all possibility of subjective bias, the criteria defining "the great champions" (performances, rewards received, etc.) will be established at the request of the CFEPP by a commission of sports specialists, notably journalists. The list will then be drawn up according to these criteria, applied in a rigorously impersonal fashion, using existing directories and sports annuals.'

Apparently this was not phrased clearly enough, as professor Ertel seems to give it a different interpretation. The journalists were supposed to set the *criteria* in so much detail or clarity that the actual selection could be done by *anybody* with access to sports dictionaries. And the criteria should be based on achievements, the protocol said.

And so it happened, more or less. The journalists advised 'take the sports dictionary of Le Roy!' Anybody in that book has accomplished great achievements in his or her sport. And at the urging of Gauquelin himself (!) the reference book *L'athlège* was added (CFEPP took the last edition, of 1951). Where the protocol essentially permitted an open ended search through various sources, CFEPP removed this potential error source of subjectivity by sticking to the two mentioned books, not chosen by themselves.

When Gauquelin finally saw the criteria he only disagreed on details. Basically he proposed to compensate for what he saw as over- and underrepresentation by adjusting the criteria for separate groups. Rugby supposedly was merely a kind of regional sport for which it was rather easy to get selected for international matches. Basque pelote (a squash-like ball game) was an international sport and had to be included. In bicycle racing all those should be thrown out who hadn't become professional. But for tennis the CFEPP criteria were too high, he wrote. He also intuitively thought there were far too many Parisians in the CFEPP's selection. In short, Gauquelin tried to introduce subjective bias into the criteria.

CFEPP's criteria

If anyone believes the sample is wrong, he should point at the criteria, and exhibit cases of false inclusions or exclusions. CFEPP had two types of criteria:

1. At least one of the two dictionaries should at least state the year of birth and a continental French town of birth unambiguously (no overseas colonies). For such a town, occurrence in at least one of three specific official manuals was required. This requirement made it possible to treat every champion in exactly the same manner. It prevented quibbles about whether the data of champions had been researched well enough or not. This was spelled out in much detail in the internal memorandum circulated by CFEPP in 1990.

2. The selected champions should be the best of France. The protocol said ‘de première grandeur’, i.e. first rate, of supreme eminence. Hence they had to be minimally either individual national champion or recordholder in their branch of sports or individually selected to represent France in an international contest (in the case of mountaineering where there are no contests, mention in both sources was required). In some sports this would exclude a lot of first class champions (notably in tennis and cycling, where there are many professionals). Now the *Dictionary* of Le Roy contains several tables of results of important annual grand competitions, such as the Tour de France, Wimbledon and the Davis Cup, and people who were mentioned in such tables as winners were also admitted.

Ertel does not say which of these criteria seem especially chosen to produce large amounts of unfair inclusions or exclusions. He does not exhibit a single champion that was unfairly in- or excluded. He probably can’t because as far as I know from communications with him, he never obtained copies of the dictionaries used. They can be consulted in Paris public libraries, which is what I did.

Ertel’s statements are untenable

Are there any outstanding champions in Gauquelin’s files, excluded by CFEPP? Suppose one uses an Ertel-type criterion: occurrence in Ertel’s database with at least two citations, not counting Le Roy citations. There are 41 such champions that were not selected by CFEPP: 4 don’t occur in the two CFEPP sources, 8 are not born in France, 7 have missing or ambiguous birth places and 5 are genuinely overlooked by CFEPP (none of these five a Mars champion according to Gauquelin).

That leaves 17 champions. Most of them were considered to be second rate by Gauquelin (!), and they were also not admitted by CFEPP for the same reason. Only one of these 17 was born in a Mars sector.

By professor Ertel’s standards, there is simply no large group of champions that is highly qualified, has a high Mars percentage, and that seems to be arbitrarily excluded by the choice of criteria. Professor Ertel only compares CFEPP’s efforts with subgroups in Gauquelin’s data and he explains differences as evidence for CFEPP bias. More about that below.

Remarkably, Ertel moreover suggests that prejudiced skeptics secretly believed there is a Mars effect, and found a way to suppress evidence for it, namely by ‘their own selection bias’ which they refused to correct after it was pointed out. Skeptics do believe things, namely that if one fairly and impartially collects data, the null hypothesis has a good chance to prevail, especially when the alternative hypothesis seems very shaky.

The missing champions

When CFEPP had selected their 1439 champion names from the sources [note 1], they wrote to the town halls. There were 98 (not 140, as Ertel states!) champions among these 1439 for whom Gauquelin had published data but for whom CFEPP did not get information. Comparison with Gauquelin’s books often learns why: the dictionary data were wrong (wrong birth date, wrong birth town, wrong first name). When I checked the whole CFEPP investigation, I found four cases where CFEPP made administrative errors, like copying the birth year wrongly or putting the forms in the wrong envelope.

In some cases the town hall provided not only the time of birth, but also the exact place of birth, or the correct birth date, either in the form of a correction or as a suppletion of missing data. The CFEPP did not copy all these extra data from forms returned from the town halls. For instance, sometimes someone of the CFEPP-workers forgot to copy the date into the CFEPP’s database. In my check of their work I have taken the position that any form returned from the town halls is to be considered as the gospel truth, unless it produces evident contradictions (one champion registered three months before his birth, and another who became high school champion at the age of 25; in the latter case the town hall must have given the data of a namesake born 9 years earlier).

This produced 52 corrections and 43 more names. Given these facts there aren’t many underqualified champions either, as Ertel tries to make us believe.

The CFEPP memorandum of 1990

The original joint protocol stipulated that a minimum of 1000 birth times of sports champions should be obtained. When the CFEPP noticed that they had passed the magical 1000 mark, they were so glad that they prepared an internal memorandum in June 1990. This contained 1071 names with complete data, among them 5 names that turned out to be not qualified (not international) at a further check. That is why Ertel speaks about 1066 champions. The 1990 memorandum was not meant for publication, because the protocol explictly forbade premature publication. The CFEPP had also neglected to sort out a small pile of town hall answers that had been kept separate from the main bulk because of some need for follow-up actions. In this pile there were 11 more names with good data [note 2]. Altogether the number of data to be added to the ones in CFEPP's memorandum was 11 + 43 = 54, thereby raising CFEPP's 1066 to 1120. The protocol was quite clear about wanting lots of data: 'The list will have to be as large as possible.'

This memorandum was shown to Suitbert Ertel in 1991 by Dutch skeptics in the naive but quite mistaken assumption that he would refrain from publishing about it until the CFEPP was ready with their final report. But professor Ertel never seems to have felt that the protocol requirements of the CFEPP test applied to him. Ertel's (1996b) views on this matter are 'Nienhuys ... did not have any qualms to add 69 (sic!) "extra" cases to the CFEPP's published database ... No protocol agreement warranted these "extra's".

Assumptions without precise knowledge

On December 27, 1994, professor Ertel asked me to clarify why the CFEPP didn't write to several birth places with names that occurred more than once in France, in other words how they came to think that only one of these was the right place. But the CFEPP rules stipulated clearly that the dictionaries should determine the birth places uniquely. Apparently he was still not quite aware of the precise criteria and didn't know that the dictionary of Le Roy almost always indicates the *Département* of French birth towns. Instead, already then he had published in several astrological and parapsychological periodicals articles about the CFEPP-investigation (Ertel, 1993b, 1993c). He used, long before the publication of the end report, data from CFEPP's internal memorandum that had been shown confidentially to him, and he phrased accusations of manipulation against CFEPP, basing them on conspiracy theories. Quite clearly he only had a dim knowledge about the study. His suggestions that CFEPP somehow had slyly avoided the 'highly eminent champions only selected by Gauquelin' goes in the same direction. How could CFEPP have known that the sportsmen and -women for whom their dictionary data (i.e. birth place or birth date) would turn out to be wrong would show such a high Mars rate in Gauquelin's publications?

Ertel's biased comments

Why didn't Gauquelin receive the list of criteria? As far as I have been able to determine, the list was sent to the person in charge of communicating with Gauquelin. When no answer from Gauquelin was sent back the selection was started. Clearly somewhere the line of communication was interrupted, and nobody noticed or checked. This is regrettable, but not as serious as Ertel suggests. Incidentally, professor Ertel has critized the U.S. 1980 test on the grounds that professor Kurtz was continually checking its progress.

Professor Ertel seems to suggest that approval of Gauquelin was necessary for the test to succeed. Not so. CFEPP was under no obligation to follow blindly Gauquelin's orders. Professor Ertel's words can be interpreted to mean that if Gauquelin had made earlier comments, they necessarily would have seemed so wise and reasonable, that they would not have been taken down and used as evidence against him, but on the other hand that their implementation would have made a lot of difference in the result. Professor Ertel even doesn't tell us what those comments might have been. [note 3]

Professor Ertel again tries to put Gauquelin's undeniable bias into a mild perspective. He has done so before (Ertel 1988), and I can only see this as an effort at cover-up. Only Koppeschaar's (1992) analysis brought out how serious this bias was, 'a futile attempt' according to Ertel (1996b).

Professor Ertel now changes his story, trying to turn the tables. The ‘impurities among Gauquelin’s sports champions’ are also caused by the skeptics! Gauquelin had been forced into these impurities, because he had to defend his position. These are all speculations on the morals of someone who is dead now. Let’s try to stick to the facts.

The fact is that in 1990 Gauquelin had at least 132 instances where he could have commented on the correctness of the data received. In my earlier paper in *Skeptiker* 1996/4 [note 4] I have explained this in more detail. In 37 cases Gauquelin had the choice to give or withhold information that might influence the result. In all 37 cases his decision pushed the result into a higher Mars rate. What about the other cases? Gauquelin seemed fond of the so-called setting and lower culmination sectors. I have not discussed this at all in my *Skeptiker* paper, but it is brought up by my esteemed opponent, so I’ll comment on it. In 12 out of 15 cases (80%) where only such a secondary sector was at stake he showed his bias. And in only 16% out of the 80 neutral remaining cases he offered corrections.

The same thing happened when he volunteered to supply missing information. Out of 39 ‘missing’ champions born in a Mars sector (according to his own records), [note 5] 37 data sets were completed by him. For the secondary sectors he was somewhat less industrious: only 73% of what he could have said was mentioned. Providing information about non-Mars champions would lower the total Mars rate in the sample of course, and here only 14% of what he knew was mentioned. His other comments show a similar pattern.

The clear bias in the Gauquelin proposals came as a genuine surprise for the CFEPP and other parties concerned. It made it necessary to make extremely careful checks on the contents of their report, causing a delay of several years in publication. [note 6]

Professor Ertel writes that CFEPP and I because of Gauquelin’s bias in his proposals, summarily reject all Gauquelin data. CFEPP comments about the reliability of earlier Gauquelin data must have escaped my attention. I myself didn’t merely assume, I took Gauquelin’s remarks as a hint for examining his earlier researches for a throw-away bias of a particular kind, and then I found rather telling evidence of him throwing away data after inspection; even though Ertel quotes my paper, he seems to have missed its main point entirely. For this purpose I laboriously examined several groups of champions, collected by Gauquelin himself (among others the ones of the Belgian Para Committee). The Mars rate among the ‘difficult to find’ was 30,9%. Among comparable champions that Gauquelin must have tried to find, but that are absent from Ertel’s files and were found by CFEPP, the Mars rate was 6,4%.

Also Arno Müller (1994) has found clear indications of earlier Gauquelin selection bias, in collections of military people and politicians. Müller tried to replicate Gauquelin’s 1971 findings with Italian writers. He found something that corresponds to my findings: the Gauquelin-only cases contained many dubious data (birth time wrong, no information from authorities) and precisely these showed a much higher Mars rate than the ones Müller found, but that weren’t found by Gauquelin. And Müller concludes that ‘there are clear indications of intentional or unintentional data manipulations.’

So we have, in order of appearance:

- evidence of selection bias from Gauquelin’s own reports about methods and his comments on the U.S. test (Kurtz et al., 1980);
- quality selection bias found by Ertel (1988);
- independent findings of Müller of several clear indications of bias (Müller, 1994);
- an astonishing and massive selection bias at a public test (Benski et al., 1996);
- evidence of throw-away bias as early as the Para test, consistent with some of Müller’s findings.

Nonetheless Professor Ertel keeps including the Para test among the ‘skeptic’s total’. Moreover, Gauquelin seems to have misrepresented facts about his own research as early as 1976 (Kurtz et al., 1997).

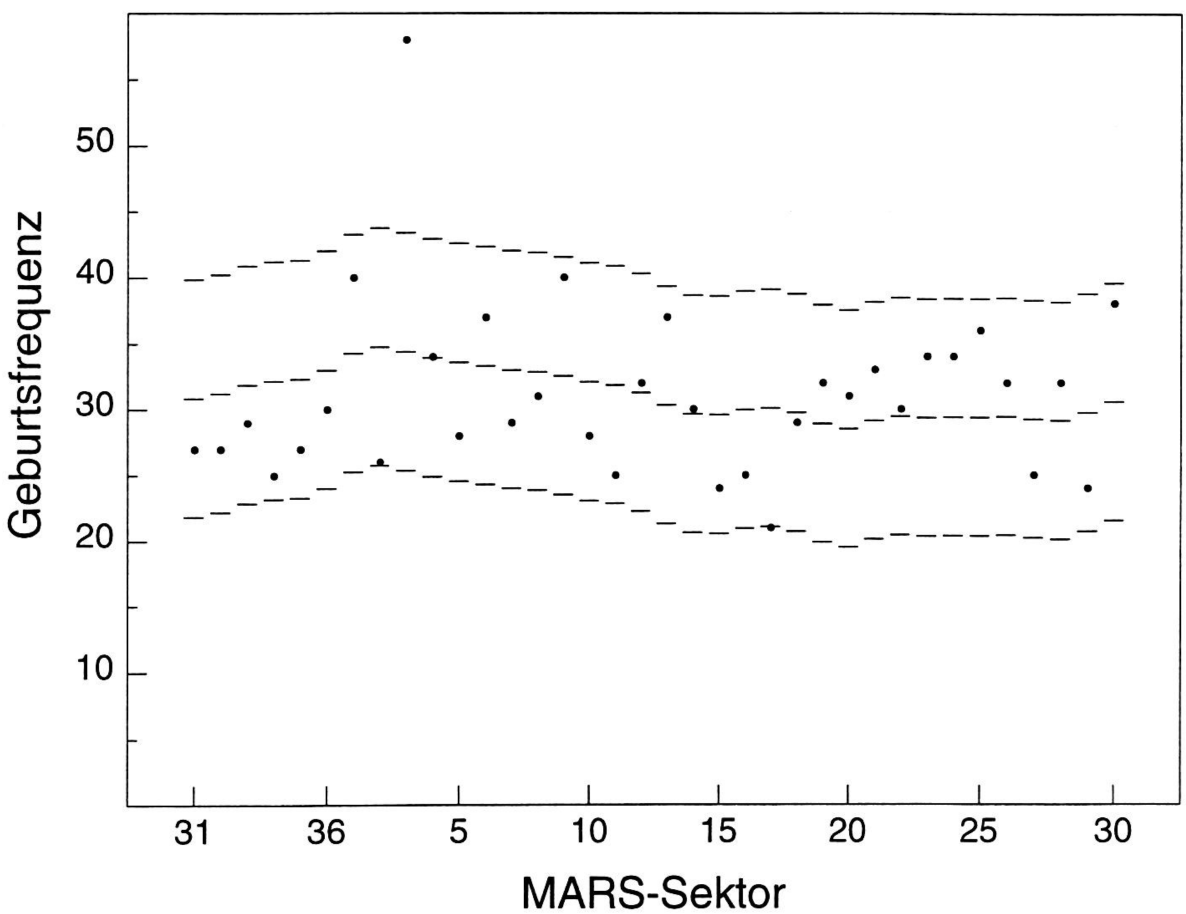
That’s enough for me. If a weak result (namely Gauquelin’s planetary effects) is completely at odds with the rest of science, admitted by Ertel (1993b, 1993c), then it should be absolutely impeccable. I dismiss arguments based on comparing data with those found by Gauquelin, including the results of the Para test. Data collection for the Para test was mostly done by Gauquelin, with the above mentioned throw-away bias as result.

Ertel's data handling

I will repeat Koppeschaar's (1992) judgment: 'No conclusions should have been drawn after [Ertel] discovered selection bias in [Gauquelin's] data.' That discovery should have been the end of the Gauquelin affair. Instead, professor Ertel tried to 'correct' for this bias by citation analysis, and then committed the error of including the very sources that Gauquelin had been using. Without these sources Ertel's eminence effect evaporates (Nienhuys 1993).

It is claimed by professor Ertel that he checked the very first example found by Gauquelin: the French physicians. In the end of 1994 Rob Nanninga noticed an inconsistency in Ertel's report (Müller and Ertel, 1994). Comparison of their tables 5 and 6 suggests that Ertel and Müller added 10 extra champions to the original 576 Gauquelin physicians of 1955, and that these 10 extra doctors all (!) were born in a key Saturn sector. And now Ertel doesn't know anymore which ones these are, and how they got there. That's Müller's responsibility, he claimed, conveniently forgetting that he wrote 'Ertel was mainly responsible for data collection' in his preface.

Statistical noise



Then professor Ertel shows a couple of graphs. To me they show just statistical noise. If the reader draws lines 9 units above and below the lower dotted line of Ertel's figure 1, then he or she will have rough 90% confidence curves. Only a few of the 36 black line data fall outside of them. In my figure here (vertical: actual number of births, horizontal sector number in the 36-sectors system)I have shown the correct CFEPP data in a such a way. The central line is obtained by randomly permuting the dates and birth times of the champions many times. Only very few data points fall outside of these intervals. This means we are dealing here with random noise that can't be interpreted.

Then in Ertel's figure 2 the 'skeptic's data' include the result of the Para test, which is highly doubtful. This dubiousness was a main result of my paper, and Ertel hasn't even tried to dispute it. Pictures that include many champions that have been de facto chosen by Gauquelin are quite misleading. And then I haven't mentioned yet the large intersection between the 'skeptic data' and the 'Gauquelin data', which is apt to produce large correlations between the two that have nothing to do with the purported 'Mars Effect'.

Hypothesis generation versus hypothesis testing

Moreover, the Gauquelin data and the CFEPP data both contain a large number of champions from *L'athlège* (about 500). These champions show a high Mars percentage (about 21%). This is not surprising. Precisely these *Athlège* data inspired Gauquelin in 1955 to think that there was a special relation between Mars and sports. Now Ertel includes these *Athlège* champions in his 'skeptic data' and in his 'Gauquelin data', so he uses the same data that served to generate the hypothesis to test the same hypothesis. It just doesn't make any sense.

`Incidentally, the CFEPP data do not show any difference in Mars rate between *Athlège* champions that also occurred in Le Roy's dictionary and the remaining *Athlèg*e champions (thus contradicting the citation related eminence hypothesis), but the high *Athlèg*e Mars rate (about 21%) was absent from the champions that were only mentioned by Le Roy (about 16%). I conclude that the high Mars rate among Gauquelin's *Athlèg*e champions was an irreproducable fluke.

CFEPP was forced to include the hypothesis forming data from the *Athlèg*e handbook, but one could say they compensated for that by taking twice as many champions as Gauquelin did in 1955. Professor Ertel's clever methods are maybe just efforts to emphasize the importance of these old champions – in other words the ones that generated the hypothesis. For example, in the 60s it was much harder to get an Olympic medal than in the 30s. The competition was much fiercer, because more countries took part. Naturally the young French champions have less medals than the old ones.

Beginning students of statistics are always warned not to draw overhasty conclusions from correlations. The example always given is the prices of Jamaica rum and the wages of (teetotalling) Methodist ministers. Behind such a relation there's merely a simple time variable that explains the rise of both. Professor Ertel fails to hide such warnings.

Gauquelin may not have included the secondary sectors in any computations, but he certainly believed they meant something and he also thought they showed some non-accidental surpluses. This is clear from his books. Why did Gauquelin show a bias during the CFEPP test with these secondary sectors? Gauquelin might have been thinking about the phrase 'Gauquelin will be free to use … the data that will be communicated to him.' What would be better than a set of data collected by card-carrying skeptic officials showing this effect as well? And how can we be certain that 1990 was the very first time this bias operated?

Ertel's table shows various small errors, but I want to point out only one of them. I recall that professor Ertel had some problem in 1992 to compute correctly p-values in case of binomially distributed variables (Ertel 1992, Koenig et al., 1992). Something similar happens here, for example the 'significant' 0.05 in column 3a should be a 'non-significant' 0.06 if the Yates' continuity correction (merely a matter of subtracting 0.5 from 297) is correctly performed, in order to get a better approximation for the exact binomial probability.

Not that it matters very much: the table shows roughly that more Gauquelin involvement makes the results more impressive. No Gauquelin involvement, namely the CFEPP test, produces for the primary sectors [note 7]:

N=1120   Mars=279   expected: 23.59% $p = 0.314$

If one includes then the setting and lower culmination sectors, one gets

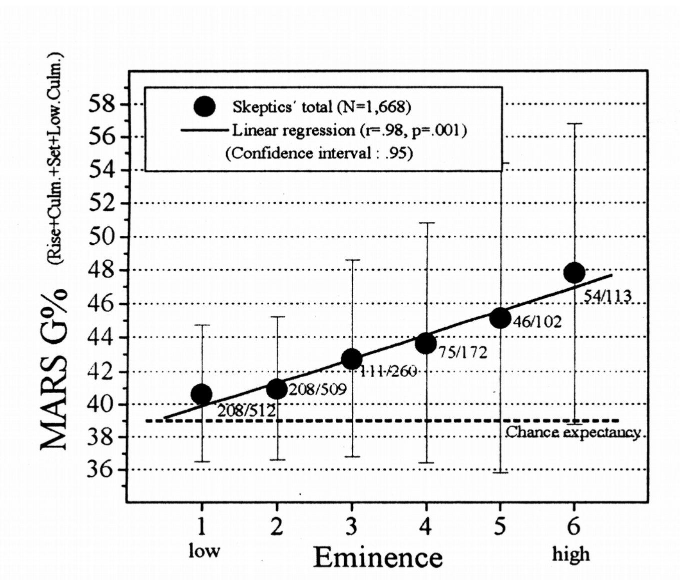
N=1120 Mars=469 expected: 39.26% $p = 0.079$

One shouldn't attach much meaning to a low value like 0.079. Not only it is not significant, but one should realize that if one performs three such tests (the original CFEPP test was also one of course), there is a 1 in 5 chance that at least one produces a *p*-value as low as 0.079.

Too beautiful

If professor Ertel wants to save Gauquelin's honor, he might emphasize how Gauquelin demolished the traditional astrology (demolished is a grand word here, because astrology continues to be practised). Gauquelin thought he had discovered an entirely new and unexpected phenomenon – the dream of each scientist. His weakness was that he used unscientific ways to explore his discovery.

Professor Ertel's results suffer from a similar weakness, and I'll explain why I think so. In his book (Ertel and Irving 1996) there is a graph that purportedly shows that the Mars percentage goes up with increasing 'eminence', i.e. occurrence in sports dictionaries. When a review in the Dutch magazine *Skepter* said the graph was wrong, professor Ertel reacted immediately, sent an 'improved' graph, and insisted strongly that his graph be published, claiming we had infringed on his copyright by displaying his first graph. I reproduce the improved graph (Ertel 1996a) below.



Each scientist has been taught early in his or her career to check data in various ways. One check is looking at the spread in the data. A huge spread in data where the measurement process predicts a small spread is evidence of an unknown error source, but a spread that's too small indicates methodological problems. I recall the fact that the unreliability of Jacques Benveniste first showed up by the fact that his data were too 'beautiful'. [note 8] Similarly, this graph of professor Ertel is too beautiful: the black dots lie too close to the line.

The fractions constitute the data. For example 208/512 at 'Eminence 1' means 512 champions had 0 citations, and 208 of these were born in a Mars sector (the numbers are not quite correct, because Ertel's computer apparently can't count well). Ertel argues that the primary and secondary sectors *together* show a very clear 'eminence effect'. The oblique line is produced by weighing these six groups with equal weights, which seems wrong to me. If they are given their proper weights, we get a slightly different line, with equation

percentage equals 40.10459 plus 1.31277 times number of citations

I give the numbers, in case a reader would like to check my computation. The slope is not significantly different from 0. Its standard deviation is 0.80614 percent per citation, yielding p = 5.17%, one-tailed (in the computation I have assumed that the total number of Mars champions is given, namely 702; also the total numbers in each class, 512, 509 and so on, I have considered as fixed). However, when we compute whether the deviations of the dots from this 'prediction' are more or less what chance says it should be, we find a chi-squared (4df) value of only 0.1950, which means a chance of only 0.46% of accidentally finding such a close match between the dots and the line, even if the trend was factual. [note 9] There are certainly not a few scientists who would voice a suspicion of conscious or unconscious data manipulation on this evidence.

Indications for data manipulation?

As I have said, this suspiciously close match is visible without computation. The error bars are about two standard deviations, and one would expect the distances of the dots to the line to be in the order of one standard deviation, and about two of them to lie even farther away than one standard deviation.

This is the explanation for the extremely high correlation coefficient shown in the legenda of the graph. To me, the graph shows two things, as long as Ertel doesn't produce plausible explanations for it: on the one hand one cannot summarily reject the suspicion of a kind of data manipulation (that may very well have been unintentional) when such clear indications exist, and on the other hand the graph constitutes a rather impressive testimony to the lack of statistical understanding of its author (the idea to compute a correlation coefficient here is dubious in itself, as these black dots cannot be thought to have a random eminence class assigned to them).

Are we seeing here the culmination of many years of selecting arguments, sources and data, ever searching for the best 'evidence' for some preconceived idea, a kind of data guided optimisation process? At any rate, Ertel's purportedly fixed 18 sources were changed when it came to analysing the CFEPP data, 60% of the citations came from four new books that replaced six old ones, and not all names were taken from the new sources either.

Let us examine whether similar curiosities occur in earlier Ertel publications. Professor Ertel (1995a) has commented on Müller's (1994) paper. Ertel agreed that there was evidence for a Gauquelin bias in collecting Italian writers (who supposedly show a Moon effect), and even adduced more evidence himself. But Ertel asked: 'Does this throw a Moon effect in writers out of the race?' and he answers this question about three pages further on: 'Gauquelin's findings were exaggerated, but fundamentally he was right.' To support his view he constructed a new eminence grading out of a combination of citation frequency and writer's achievements (number of books written). The result is displayed in the same kind of diagram as what he sent to *Skepter*. My analysis shows for this diagram the same: the trend is nonsignificant (p = 0.11, one-tailed), but the dots lie so close to the line that one would expect this to happen by chance only in 1.5% of such analyses, whether the trend is real or not. Even so, professor Ertel remarks that the trend in this graph is not altogether monotonic.

The same paper shows more such diagrams, among which one for the published Gauquelin champions. The same computation, again for six eminence classes, shows a significant result there (p = 0.015, one-tailed) but again the match of the dots to the approximating line is suspiciously close: the odds against chance are 3,4% in this case.

My contribution to *Skeptiker* has explained how a Gauquelin throw-away bias could be related to the number of sources he had available. The observed trend – insofar as it is real – might be a side effect of this Gauquelin bias, hence it might not be entirely a product of Ertel's data handling procedures. [note 10] However, these repeated close matches in Ertel's publications point to methods of defining 'eminence' that can hardly be called 'blind'.

I would have hesitated to point out all of this, if professor Ertel himself wasn't in the habit of pulling out one statistic after the other after the tests are over and insinuating improper methods or evidence of yet another transcendental truth.

Meaningless discussions

I do not see a point in discussing the Mars Effect further with professor Ertel. I'll explain why I take that position. In the fall of 1992 professor Ertel was creating 'random' data by shifting the birth times of 1076 champions by multiples of one year. At that time I merely spent my lunch breaks debating him on the Internet. I pointed out that the Mars positions in the sky repeat almost exactly every 32 years and somewhat less exactly in 15 and 17 years and that hence obtaining 51 yearly shifts is meaningless for statistical experiments. Moreover this 32 year period seemed inexplicably absent from his data. Something seemed wrong. He then switched subjects, and I dropped the matter, having no access to any original data.

After publication of *The "Mars Effect"* (Benski et al., 1996) two things happened. Professor Ertel broadcast a long article on the Internet, titled 'On Dr. Nienhuys' stamping the evidence ...'. In this he alleged (on the basis of sloppy reading) that I had omitted 222 champions. Of course I did no such thing. I gave detailed comments about the CFEPP research, and meticulously accounted for problem cases almost to the individual level. I'm unaware of professor Ertel ever retracting this false allegation.

Later he sent me precisely the same data about yearly shifts again (in graph form), and I explained again, referring to my November 1992 remarks. I also observed that it looked as if somehow a 'shift by one or two years back' was mislaid and inserted between 'shift 18 years forward' and 'shift 19 years forward'. [note 11] Towards the end of 1996 (Ertel 1996b) he repeated this 1992 argument in print, accompanied by the remark '[Nienhuys] resorted to confusing statistical constructions even though a look at [the picture] tells it all without any statistics.'

In the same paper he falsely alleged that a forthcoming paper by Paul Kurtz, me and Ranjit Sandhu (Kurtz et al., 1997) contained certain 'arrogant and appalling' quotes. [note 12] Moreover, as soon as I saw the overly beautiful graph he had submitted to *Skepter*, I let it be known, in no uncertain terms, what I thought of it. He merely surmised that I probably had made a small error.

That's why I gave up trying to argue basic science with him: he doesn't understand and he doesn't listen either, but spreads false allegations. In private communications he uses various threats [note 11], he publishes about matters that have been shown to him confidentially and quotes without seeking approval remarks made in private letters. That's why I refuse to have any direct communication with him. I'll join Jean-Paul Krivine, secretary of CFEPP, in saying: 'It's over.'

Notes added 2017:

1. These 1439 had been selected already in 1986, two years before the CFEPP was able to send out requests to town halls.

2. When I found those last town hall reactions they were in a light green folder marked 'Have these been sorted out or not?' Most of these answers were dated in July 1989, about a year before the Memorandum, the others were not dated. Possibly coding the results (at another location than the Paris office of CFEPP) had started already before these letters arrived. Anyway, there was no written statement, for instance that town hall returns after a certain date should be ignored.

3. If Gauquelin had seen the original criteria and the list of 1439 selected persons, he might have pointed out quite a few selection errors. Supposing that would have happened, then it is not unreasonable to think that CFEPP would have checked the *entire* list once more and then they would have found the five percent selection errors I found.

4. I.e. the translation of the article in *Skeptical Inquirer*.

5. Ertel had visited Gauquelin well before 1990, and not merely did a perfunctory check on his archives, but acquired many data of Gauquelin, specifically also his unpublished ones. He then added information about various handbooks in which these names occurred and published about it

in 1988. Ertels computer files (names, birthdates, Mars sector, sources, but no birth times) about sports champions were given to the Dutch skeptics when they wanted to investigate a possible explanation of the Mars effect in 1991.

6. No one in CFEPP seemed to be aware of Ertel's 1988 paper, let alone that they bothered to look into Gauquelin's 1970 publication on his 2088 champions or into Ertel's files. This explains their surprise. But when Kurtz decided (1993) to publish an English translation of CFEPP's report, the seriousness of Gauquelin's attempted fraud made it necessary that the report was thoroughly checked, and Kurtz asked me to do that.

7. Gauquelin used various types of sectors. Originally there were 12 sectors, and then sectors 1 and 4 were the ones that mattered. But at some time he thought a division into 36 sectors was better. The sectors 'that mattered' were then 1,2,3 and 10,11,12 in the 36-sector system. However, he arrived at the conviction that to this should be added 36 and 9, slightly enlarging the original choice, meaning that those born before just before Mars rise or just before Mars culmination also counted. So the 'primary sectors' were 36, 1, 2, 3, 9, 10, 11, 12. The secondary sectors were 18, 19, 20, 21, 27, 28, 29, 30.

8. Benveniste gave in the famous *Nature* homeopathy paper a table that claimed to contain many means of triple measurements. But the stated standard errors of these means were far below what must be expected for that kind of measurement.

9. The computation is not entirely correct. The value of the slope is a weighted sum of (here) the Mars-numbers in the different classes. The proper model should take into account that their sum is given, so the Mars-numbers follow a hypergeometric distribution. Many handbooks of statistics give formulas for their variance and covariance (which I discovered only after deriving them myself). The variance is $(N - K)/(N - 1)$ times the variance in case these Mars-numbers had followed individual binomial distributions; here N is the total number 1668 of champions, and K is the total number in the relevant eminence class. The computation of the probability of all the points being so precisely on a line then gives a chi-squared of 0.19517, rather than 0.1950 as originally stated. A simulation of the process is as follows. A total of 1668 balls, 702 of them red and the remainder white, is randomly distributed over 6 boxes in such a way that they contain 512, 509, 260 etc. balls. In each simulation the best linear approximation of the red ball nummbers is constructed and S, the weighted sum of squares of deviations is noted. Of 10,000 experiments only 53 gave an S as small as Ertel's.

Then in a roughly similar way it was ensured that the probabilities of a red ball landing in the separate boxes was 205/512, 211/509, 111/260 etc. (as close as possible to Ertel's trend), still keeping the total number of red balls equal to 702. After 100,000 such experiments 476 had yielded an S as small as Ertel's.

So even though my original calculation wasn't entirely correct, the figure 0.46% is quite correct and backed up by various simulations.

One more remark: the black dots do not come from a randomly distributed population. However, if each of the 1668 sportspeople is considered as a separate dot the eminence class can be considered somewhat as a random variable. If one thinks of the total sample as split into 1668 non-empty classes, one obtains $p = 0.4999$. Standard statistical programs produce a similar result.

10. The conjectured throw-away bias of Gauquelin consisted of retaining or throwing away data that seemed unreliable, and basing the decision on the Mars sector. If there was more than one source for data, the probability is higher that their data contradict each other, thus raising doubt.

11. Details of my conjecture about Ertel's error are omitted in German translation.

12. These unpleasant remarks were omitted at the request of the editors of *JSE*, who sent us (Kurtz, Sandhu and me) a very long referee report obviously written by Ertel. Ertel assumed without any reason that his objections were altogether ignored.

13. For example, I went to Paris in November 1994. Ertel knew that and asked me specifically for some information about certain champions mentioned in the CFEPP Memorandum. When I returned I found an e-mail from Ertel (Nov. 25, 1994) about the right to publish Gauquelin's letters (the ones with the proposed corrections). "If he were alive … he might defend himself by a legal suit against the CFEPP's publishing his letters … I will inform Michel's sister and his son about the CFEPP's intentions, perhaps they are able and want to defend the rights of the deceased." A bit later he started threatening again to publicize all our private correspondence, if I didn't do exactly as he wanted. That Ertel wouldn't hesitate legal action in scientific disputes became apparent three years afterwards. At that moment I decided that I didn't want to correspond anymore with him. I changed my mind later after being urged by a Dutch skeptic colleague to cooperate a little better with this serious researcher. I still regret it bitterly.

References

Benski, C. et al. (1996) *The "Mars Effect"*. Prometheus, Amherst.

Ertel, S. (1988) Raising the hurdle for the athletes' Mars Effect: Association co-varies with eminence. *Journal of Scientific Exploration* 2 (1), p. 53-82.

Ertel, S. (1992) Reanalyse des Kasseler Wünschelruten-Tests der GWUP. *Skeptiker* 5 (3), p. 69-72.

Ertel, S. (1993a) Puzzling Eminence Effects Might Make Good Sense. *Journal of Scientific Exploration*, 7 (2), p. 145-154.

Ertel, S. (1993b) Die Gauquelin-Planeteneffekte: Forschungsstand 1993. *Grenzgebiete der Wissenschaft* 42, p. 99-114.

Ertel, S. (1993c) Die Gauquelin-Planeteneffekte: Forschungsstand 1993. *Meridian* 5, p. 6-12.

Ertel, S. (1995a) Die Stärke des Gauquelin-Planeteneffekts: Arno Müllers Bilanz korrekturbedürftig. *Zeitschr. f. Parapsychologie und Grenzgebiete der Psychologie*, 37 (1/2), p. 3-27.

Ertel, S. (1995b) Reply to Dr. Nienhuys' letter. *Correlation* 14 (1), p. 50-54.

Ertel, S. (1996a) *Het weerbarstige Marseffect*. (letter) Skepter 9 (2), p. 43.

Ertel, S. (1996b) How to Suppress the Gauquelin's Mars Effect? Strategies of Concerned Committees. *Correlation* 15 (1), p. 2-16.

Ertel, S. and I. Irving (1996c) *The Tenacious Mars Effect*. Urania, London.

Koppeschaar, C.E. (1992) 'The Mars effect unriddled'. In: *Science or Pseudo? The Mars effect and other claims. Proceedings of the Third EuroSkeptics Congress, October 4-5, 1991, Amsterdam*. Skeptische Notities 8, J.W. Nienhuys (ed.), Stichting Skepsis, Utrecht.

Kurtz, P., J.W. Nienhuys, R. Sandhu (1997) Is the "Mars Effect" Genuine? *Journal of Scientific Exploration*, 11 (1), p. 19-39.

Kurtz, P., et al. (1980) Four-Part Report on claimed ‘Mars Effect’, *Skeptical Inquirer* 4 (2) p.19-63.

Müller, A. (1994) Gauquelin-Effekt: Eine kritische Bilanz. *Zeitsch. f. Parapsychologie und Grenzgebiete der Psychologie* 36 (3/4) p. 131-162.

Müller, A., and S. Ertel (1994) 1083 members of the French “Académie de Médecine”. A.P. Müller Verlag, Waldmohr.

Nienhuys, J.W. (1993) Comments on Puzzling Eminence Effects. *Journal of Scientific Exploration* 7 , p. 155-159.

Nienhuys, J.W. (1996) Wie der Mars-Effekt zustande kam. *Skeptiker* 9 (4), p. 124-127.