Journal of Scientific Exploration, Vol. 7, No. 3, pp. 283–292, 1993 0892-3310/93

Comments on Dutch Investigations of the Gauquelin Mars Effect

SUITBERT ERTEL

Institut fur Psychologie, Gosslerstraße 14, 3400 Goettingen, Germany

Abstract — The first of two Dutch skeptics' attempts at disproving the Mars effect failed (Nienhuys 1993a). Contrary to the Nienhuys view, however, the second pass of the Dutch critics at the Gauquelin claim — an attempt at undermining the validity of his data base — is shown to fail as well. The critics drew apparent support from my previous unearthing of a Gauquelin bias— Gauquelin had exempted cases from publication (Ertel, 1988). Yet they neglected the fact that any such bias had been neutralized as a result of my pooling of published and unpublished data. Specifically, a significant eminence trend was demonstrated in the unmanufactured total sample in my 1988 report. In addition, Dutch endeavours at rendering the eminence relationship insignificant either failed (even a less sensitive scale with 12 instead of 36 sector division yielded significance) or were illegitimate (splitting up of the entire sample for that purpose violates methodological logic). Thus, the present (fourth) attempt in the history of resistance against the Gauquelin challenge by organized skeptics has added two misses to their record.

Introduction

The Gauquelin Mars effect has aroused a good deal of opposition and corresponding attempts at empirical refutation by four skeptic committees. The first was the Belgian attempt, which failed (Comite Para, 1976). Another negative result was obtained in the first of two U.S. American efforts ("Zelen test," Zelen, Kurtz, & Abell, 1977). The results of the second U.S. effort (U.S. athletes test, Kurtz, Zelen, & Abell, 1979 — 1980) did not unequivocally contradict Gauquelin's claim either (Ertel, 1992b). A recent French replication, based on new data (Benski, 1991) is now being studied — (Ertel, in preparation). The most recent attempts at refutation have been investigations by Dutch skeptics, published in The Proceedings of the Third Euroskeptics Congress (de Jager & Jager, 1992, Jongbloet, 1992, Koppeschaar, 1992, Nienhuys, 1992; rejoinders by Ertel, 1992a, and Schneider — Gauquelin, 1992), and accounted for by Dr. Nienhuys in the present issue of JSE (Nienhuys, 1993a), henceforth "N93."

Dutch Attempts I and II

The Dutch approach consisted of two attempts, the first of which searched for an alternate explanation of the Mars anomaly, an issue that was expounded at the Euroskeptics Congress. It did not succeed as shown in my critique (1992a) (1) and as admitted by de Jager & Jager (1992) as well as by Nienhuys ("essentially their findings were negative," N93).

The second attempt consists of questioning the soundness of the Gauquelin data base and includes a public relations campaign (2): The Mars effect was turned into an effect of biased sampling, explicitly by de Jager: "this [our] finding killed the reality of the [Mars] effect" (personal communication, letter Dec 10, 1992); less explicitly by Nienhuys, who merely cast doubt on the data leaving conclusions to the reader: "the Gauquelin data suffer from bias." In what follows I will argue that attempt II of the Dutch refutation of the Mars effect also fails.

Attempt II of the Dutch investigations did not come up with new facts, since support for the negative conclusions was drawn from my JSE 1988 account of Gauquelin athletes data — specifically, my comparative data analysis of published vs. unpublished data, which provided evidence for a Gauquelin selection bias.(3). Nienhuys' interpretation of what I found and of additional Dutch observations, however, are not in line with my own, so the divergence demands scrutiny.

Pooling Data Repairs Doubtful Subdivisions

As early as 1955, Michel Gauquelin in his pioneering study on French professionals subdivided three of nine samples using the criterion of "eminence" (Gauquelin,1955). Eminence in this case was operationally defined by lines devoted to individuals in biographical dictionaries (painters: high-medium-low eminence, p. 233), or by numbers of expert nominations (actors: high-low eminence, p.234; politicians: high-low eminence, p. 235). More pronounced planetary effects were reported for the more eminent as compared to less well-known subsamples. In his classic replication study of 1960 on other European samples (Italian, German, Belgian, Dutch), Gauquelin allegedly found support for his eminence claim for all professional groups (N = 13, Gauquelin, 1960, p.160ff). Operational definitions of eminence, however, had not improved — on the contrary, subdivisions of samples had largely been based on sloppy criteria. Among athletes, for example, Gauquelin considered all Italian football players (N = 599) and all German athletes (N= 118) except 8 (pp. 85, 89) as not reaching the criterion (more instances are given in N93). In one major publication of sports champions data (Gauquelin & Gauquelin, 1970) mediocre figures were not entered. This gave sufficient grounds for my travelling to Paris in 1986 with the intent to search in the Gauquelin archive for missing data. I came up with a great number of unpublished athletes' birth dates (N = 1,503), see detailed description in Ertel (1988).

No doubt, Gauquelin's repeated eminence observations had entitled him to select for publication successful samples and to discard mediocre ones. Selections, however, should have been done with utmost objectivity as well as replicability. In this regard Gauquelin, generally a model of painstaking

TABLE 1
Citation percentages for published and unpublished Gauquelin athletes.

Citation frequencies	Published (N = 2,888)	Unpublished (N = 1,503)	Difference
0	46.1	62.5	16.4
1	-22.8	29.4	10.5
2	16.0	4.9	-11.1
3	7.3	2.9	-4.4
4	3.3	0.2	-3.1
5	2.7	0.0	-2.7
6	1.1	0.0	-1.1
7	0.6	0.0	-0.6
8	0.1	0.0	-0.14
	100.0	100.0	

methodological care, had violated norms. I conclude that this had occurred inadvertently rather than on purpose since Gauquelin facilitated my search for unpublished data without reservation.

My pooling of all athletes data and my subjecting them to citation counts repaired past sampling flaws. Thus, a subsequent reanalysis of the entire data pool showed first that birth dates not published by Gauquelin were in fact predominantly those of least-cited sports people (see Table 1). Secondly, I also found that Gauquelin's discarding of athletes, in a subsample of N=659, had been affected by his being aware of Mars sector positions. Any such mismanagement of data, however, including biased selection, was now undone by pooling the data. Pristine conditions, as it were, were thus restored for testing afresh Gauquelin's contention. The eminence trend as shown in Figure 3 of my 1988 paper (p.72), was the result of an unhazarded new approach.

Refutation of Nienhuys' Arguments

The Dutch critique failed because it neglected the improved methodological ground work just described. The first of Nienhuys' arguments is merely a paraphrase of what I had already published and is thus redundant.

Nienhuys' second argument addressing my eminence result ("Koppeschaar finds no eminence effect") deserves three explanatory comments. First, in Figure 9 of his critique Koppeschaar presented Mars key sector proportions of the total of 4,391 athletes, separately for 9 citation ranks (4). His conclusion regarding trend ("no significantly increasing eminence effect is evident"), however, was based on visual inspection only, which is deceptive. That is, Kendall's tau-statistic applied to his numbers reveals a significant eminence trend ($p = .025$). I provided Koppeschaar the result of my calculation and offered my trend program for him to check this, but he did not make use of this offer.

Second, Koppeschaar's calculations utilized Gauquelin's former 12 sector division (two of 12 sectors are defined as key sectors, i.e., sectors 1 and 4 on that scale) instead of the now imperative 36 sector division (8 of 36 sectors are defined as key sectors: nos. 36, 1, 2, 3, 9, 10, 11, 12, on that scale). The 36 sector definition applies to all professions (not only athletes) and to all planets (not only Mars), i.e., it does not just favor the present study. The Dutch skeptics' indifference to scale requirements (Nienhuys excludes this issue as well) despite my pointing the problem out in personal correspondence (Koppeschaar — Ertel), should be kept in mind. The error probability of an eminence trend for athletes' Mars sector proportions, when based on 36 sector precision, is raised from $p = .025$ (12 sectors) to $p < 0.005$ (36 sectors).

Thirdly, Koppeschaar's and Nienhuys' subdividing of the total sample into 4 (or whatever) subsamples and then subjecting them to separate significance tests is not legitimate. This Machiavellian "divide and rule" strategy (i.e., breaking the sample down to get rid of degrees of freedom, thus slipping away from statistical confirmation) had its predecessors in the Zelen test debate and was criticized (Tarkington, 1980, p.19, Curry, 1982, p.37) (6). It is well known that dividing a total sample up into subsamples may be useful for testing the homogeneity of that sample. Thus, the key sector (KS) percentage of, say, a French subsample may be compared, using a Chi-square test, with KS percentage of the rest (non-French) of the total sample. If the French sample's KS percentage level differs significantly from the level of the non-French total a prematurely generalized hypothesis (the Mars effect applies for all nationalities) might possibly be challenged(7). But the Dutch strategy was different. Subsamples were not compared with the total sample, but were instead subjected to separate significance tests (8).

Nienhuys' third point ("Statistical checks of Koppeschaar's argument") was based on a trend test proposed by Nienhuys and tested by Dijkstra. This test served as an alternative to more familiar published procedures. It was applied, as it should be (assuming the details are correct), to the total of N = 4,391 athletes. It ought not to have been applied, however, as was shown above, to 12-sector data. Nevertheless, despite the exclusion of 25 percent of the pertinent information, the trend test proved significant ($p = 0.033$) which is close to the $p = 0.025$ level as obtained for 12-sector data by Kendall's tau (see above). But Nienhuys' main argument does not bear on this result at all, dwelling instead on nonsignificant eminence results for separate subsamples, above all for the French.

Nienhuys and Dijkstra's trend test was also applied separately on published and unpublished subsamples, but that approach was unjustified. Gauquelin had separated (and not published), for good reasons, mediocre sports people from more successful ones (see Table 1 above). Separate correlations with eminence for published and unpublished samples, as performed by Nienhuys and Dijkstra are thus bound to decline. As an illustrative analogy consider that the correlation of body size between husbands and wives in a population must decline if done separately for samples of tall and short husbands.

Nienhuys' fourth argument with its minute account of Gauquelin's selecting successful samples does unquestionably point out some flaws, but flaws which I had suspected in 1986 and which had been corrected in 1988 by my pooling of all available data.

Nienhuys' last argument is his simulation. However this caricature of how deliberate and biased selections of good achievers might proceed does not add any new aspect to the question the Dutch desired to settle by negation: "Is a Mars effect present in the Gauquelin athletes' data?" De Jager apparently did not go into the matter deeply enough, even mixing up a crucial sampling division (cf. Nienhuys' endnote), thus missing entirely the logic of my 1988 eminence result. For him as an astronomer the "Mars effect" issue apparently did not deserve serious consideration (9). It is more difficult to excuse Nienhuys' for neglecting the main issue. An extensive correspondence about the Mars effect on two e-mail forums preceded the present publication (10).

The Crucial Issue

Only Koppeschaar, after having been urged in my letters not to neglect in his paper the crucial point of my 1988 article, commented on the essence of what I had conveyed. He quoted from my paper:

"The presence of selection bias, therefore, did not weaken the conclusion that Mars' positions and the athletes' births are statistically related. Correction for selection bias by pooling all records increased empirical support for the stronger version of this claim [= for the eminence claim]: The data have overcome, in spite of disturbing effects of bias, the higher methodological hurdle" (Koppeschaar, p.183).

Koppeschaar, however, took issue with this contention: "I do not agree with this conclusion. How can any trend analysis be based on selection-biased data?" My reply: Selection bias had been removed by pooling published and unpublished data. Any effect of previous subdivisions had thus been rendered ineffective.

Koppeschaar continued: "Especially if one closely inspects the absence of higher ranks of eminence for the unpublished data, it is immediately clear that a trend analysis for the total sample is greatly influenced by the outcome of the already published sample."

My comment on that: Koppeschaar seems to have noticed eventually, first, that Gauquelin's unpublished data are associated with lower eminence ranks as compared to his published data; second, that Mars key sector percentages are lower for unpublished than for published data; and third, that pooling both samples should therefore give rise to an upwards eminence slope just as pooling tall and short husbands should increase body-size correlations with their wives above the level of separate correlations.

Koppeschaar, however, distrusts this compelling logic. In view of the astounding difficulty my Dutch opponents have met with at grasping the core of my message of 1988 it may be helpful to repeat that the eminence slope of athletes (N = 4,391) is unrelated to the ways Gauquelin managed or mismanaged previous data subdivisions. That is, the eminence vs. Mars-births relationship would have taken exactly the same shape even if entirely different decisions had been made by Gauquelin regarding which data were or were not to be published (11).

Future Perspectives

Two great advantages are associated with eminence trend assessment in the Gauquelin research: First, such trends are independent of key sector proportions. In other words, significant trends can exist and thus reflect planetary connections even when the average key sector level across eminence ranks is low and thus lacks such a connection. Second, trend tests do not require expectancy calculations whose rationale, as the Comite Para demonstrated, might remain a matter of debate. That is, the trend test approach to the Mars effect hypotheses circumvents one of the greatest past obstacles impeding agreement among researchers. The present debate has cast light upon improved conditions for joint future endeavours in this field.

Postscript

Seven brief remarks on Dr. Nienhuys' comments (Nienhuys, 1993a) on my Munich paper (Ertel, 1993):

(1) Dr. Nienhuys' reanalysis of my eminence calculations confirmed the trend i.e., Mars key sector births increase with increasing eminence as shown in his Figures 1 and 3. The Mars effect itself is thus confirmed.

(2) Nienhuys, however, says that when citations from three sources are deleted, this trend disappears, as shown in his Figure 2. The three sources deleted here are important in that they are among the few covering most sports fields; they exceed others in regard to eminence information as becomes evident in Nienhuys' Figure 3.

(3) The unexplained "modus operandi of the Gauquelin bias" (Nienhuys) seems to imply that the mere fact that these books were borrowed from Gauquelin introduced bias. The critic does not indicate how bias might have worked here and did not provide any reasonable answer to my explicit question in personal correspondence.

(4) In order to bring the trend curve down in Figure 2 Nienhuys not only deleted citation sources, he also switched to the less sensitive of the two sector definitions (2/12 instead of 8/36 sectors). He knew from Koppeschaar that by switching to the 2/12 key sector definition the eminence trend would decline. Nevertheless he combined, without informing the reader explicitly, both trend-lowering contributions in his Figure 2, whose difference from Figure 1 is therefore ambiguous.

(5) Nienhuys refers to my discussion of eminence relationships as an "exercise" in arbitrary curve-fitting. No justification is given for this and my requests in subsequent correspondence to point at specific errors or statistically illegitimate decisions were not answered.

(6) Nienhuys' conjectures in his last paragraph became comprehensible only after having requested and received from him additional explanations. His intent was to explain downward key sector trends by selection bias. On noticing that his bias idea is not supported by the data (and not noticing that his idea is contradicted by the observed upwards trends anyway) he says: "Therefore I think that selection by ... [that fancy bias] cannot be the only source of the Gauquelin bias, other methodological errors may also be responsible" (p. 159). With this twist his whimsical construction is maintained despite counter evidence; the error he himself committed in inventing it is converted into a Gauquelin error, and the explanation for downwards and upwards eminence trends which he failed to provide is expected from more errors of such kind in the future.

(7) Dr. Nienhuys said at the beginning: "It is not easy to comment on a report in which no data are presented that permit any form of numerical analysis." The reader is not told (respective acknowledgments are missing) that I provided all the data that he used for his Comments and that he did not ask for what he might have wanted in addition.

References

Benski, C. (1991). Verification de l'effet de Mars. Etat de I'experience au 20 Juin 1990 (Mineographed report). Comite Belge pour l'Etude des Phenomenes Reputes Paranormaux (1976). Considerations critiques sur une recherche faite par M. M. Gauquelin dans le domaine des influences planetaires. Nouvelles Breves, 43, 327-343.

Curry, P. (1982). Research on the Mars effect. *Zetetic Scholar,* 9, 34-53. de Jager, C., & Jager, R. (1992). Spurious periodicities in planetary correlations and the Mars effect. In J. W. Nienhuys (Ed.), Science or Pseudo? The Mars effect and other claims (Proceedings of the Third EuroSkeptics Congress. October 4-5, 1991, Amsterdam, pp. 129-147). Utrecht: Skepsis.

Ertel, S. (1988), *Journal of Scientific Exploration,* 2, 53.

Ertel, S. (1992a). Mars effect survives critique of Dutch skeptics. In J. W. Nienhuys (Ed.), Science or Pseudo? The Mars effect and other claims (Proceedings of the Third EuroSkeptics Congress. October 4-5, 1991, Amsterdam, pp. 185-204). Utrecht: Skepsis.

Ertel, S. (1992b). Update on the Mars effect. *The Skeptical Inquirer,* 16 (2, Winter 1992), 150-160.

Ertel, S. (1993). Puzzling eminence effects might make good sense. *Journal of Scientific Exploration, 7,* 145

Ertel, S. (in preparation). *Tenacious Mars effect: scrutinizing data from skeptics files: CP, CSICOP, CFEPP.*

Gauquelin, M., & Gauquelin, F. (1970). Sports champions, 1 - 2.089. *Series A: Professional notabilities* (Vol.1). Paris: LERRCP.

Gauquelin, M. (1955). *L'Influence des Astres,.* Paris: Dauphin.

Gauquelin, M. (1960). *Les Hommes et Les Astres.* Paris:Denoel.

Jongbloet, P. H. (1992). Circannual and circadian biological rhythms in relation to eminence. In J. W. Nienhuys (Ed.), Science or Pseudo? The Mars effect and other claims (Proceedings of the Third EuroSkeptics Congress. October 4-5, 1991, Amsterdam, pp. 148-161). Utrecht: Skepsis.

Koppeschaar, C. E. (1992). The Mars effect unriddled. In J. W. Nienhuys (Ed.), Science or pseudo? The Mars effect and other claims (Proceedings of the Third EuroSkeptics Congress. October 4-5, 1991, Amsterdam, pp. 162-184). Utrecht: Skepsis.

Kurtz, P., Zelen, M., & Abell, G. 0 . (1979-1980). Results of the U.S. test of the "Mars effect" are negative. *The Skeptical Inquirer,* 4(2), 19-26.

Nienhuys, J. W. (1992). Introduction to the Mars effect. In J. W. Nienhuys (Ed.), Science or Pseudo? (Proceedings of the Third EuroSkeptics Congress. October 4-5, 1991, Amsterdam, pp. 112-128). Utrecht: Skepsis.

Nienhuys, J. W. (1993a). Comments on puzzling eminence effects. *Journal of Scientific Exploration,* 7, 155

Nienhuys, J. W. (1993b). Dutch investigations on the Gauquelin effect. Journal of Scientific Exploration,7, this issue.
Schilling, G. (Ed.). (1991). Het laatste heilige huisje van de astrologie stort in. de Volkskrant,Oct. 5.
Schneider-Gauquelin, F. (1992). Examining rational explanations of the Mars effect. In J. W. Nienhuys (Ed.), Science or Pseudo? The Mars effect and other claims (Proceedings of the Third EuroSkeptics Congress. October 4-5, 1991, Amsterdam, pp. 204-223). Utrecht: Skepsis.
Tarkington, E. (1981). Gauquelin's travels. Phenomena, 2(2), 18-20.
Zelen, M., Kurtz, P., & Abell, G. O. (1977). Is there a Mars effect? The Humanist, 37, 36-39.

Footnotes

1. My critique was based on new analyses of Gauquelin data that put the Dutch claims to the test. Mme. Schneider-Gauquelin (1992) drew attention to already available counter-evidence from past Gauquelin research.
2. "Campaign" alludes here to the Dutch authors' seeking publicity for their first approach by informing the national newspaper "Volkskrant" which headlined a corresponding article "The last citadel of astrologers is crashing" (Schilling, G., 1991). De Jager, Jongbloet, and Koppeschaar proclaimed here in public successful demolition of Gauquelin's basic claim.
3. I drew the skeptics' attention to my 1988 paper. This occurred in my letter to Koppeschaar, Jan. 30, 1992, who apparently had not taken notice of that publication. De Jager was notified correspondingly by members of his group.
4. Koppeschaar refrained from putting citation frequencies 6-9 into one rank =>6 (as I had done in 1988) since rank 8 (N= 18) and rank 9 (N = 3) did not continue the slope displayed by ranks 1 through 7. For him this was an optical advantage.
5. A copy of my trend calculation was posted to Nienhuys as well, but his second argument shows he also preferred to disregard my statistical correction of Koppeschaar's visual analysis.
6. A quote from Tarkington (1981, p.19), on the Zelen test: "Zelen et al. repeatedly broke the total sample down into samples too small to reach statistical significance. No ordinary statistical case, however strong, can survive splitting of the data into small parts. It is a fundamental concept of statistics that a small sample is much more variable then a large one and much less reliable. We can be sure that small samples will show us nothing significant in the way of proof."
7. Challenged by the Dutch skeptics' opposition I checked the internal consistency of the database using χ^2. None of the six nationalities (French, German, Italian, Belgian, Dutch, US) deviates significantly regarding Mars key sector percentage from the respective total (see Table).

The Gauquelin athletes total sample (N = 4,391) broken down regarding nationality, key sector frequencies (KS) and significance levels for deviation from total based on Chi-square (nKS = non-key sector frequencies).

Nation	FRA	ITA	BEL	GER	NET	USA
Total	2040	1328	399	154	60	351
KS	534	333	99	37	16	91
nKS	1506	995	117	300	44	260
KS %	26.2	25.1	24.8	24.0	26.7	25.9
$\varkappa^2$	.52	.35	.17	.23	.03	.01
df	1	1	1	1	1	1
P	n.s.	n.s.	n.s.	n.s.	n.s.	n.s.

8. Koppeschaar correctly noticed a discrepancy regarding citation frequencies (should not be mixed up with key sector percentages) between French and non-French athletes. I was aware of this shortcoming: I had utilized the "Dictionnaire des Sports" as citation source even though this book had also served Gauquelin as retrieval source from which he had drawn a great number of cases. The reason for my decision was that biographical books such as the Dictionnaire covering many sports fields and nationalities are very rare. The gain of information provided by Dictionnaire citations regarding non-French nationalities (non-French in that book are predominantly famous) was associated with a loss of discrimination among French athletes (the French majority in this book is less famous). I might have selected from the Dictionnaire non-French citations only. But fear of crude critique by inconsiderate skeptics for "manipulating" assessment of eminence warned me to avoid more complicated decisions.
9. I quote from de Jager's letter to me, Dec. 10, 1992: "... we wrote our paper and realized this was the moment to stop searching for phantoms and to return to more promising research, which we actually did. Rieks [Jager, coauthor] returned to his X-ray sources and I to my unstable supergiants." In de Jager's view, as I see it, planets in the solar system may not compare regarding size with supergiants in distant galaxies. But if Gauquelin effects are real then planets, regarding human affairs, might prove incomparably more important.
10. The record will be provided on text file upon request. Almost all of it is also retrievable from files SKEPTIC LOG9110A-E LOG9211A-E by listserv at YORKVM1.
11. Prof. de Jager in a letter to me (Dec. 10, 1992) wrote: "I think you made a methodological error by throwing Gauquelin's initial data set (which is data set A in the above terminology) and the 2000 additional ones (B)

into one basket. That spoils both data sets, and does not allow any conclusion." De Jager does not explain why for him throwing samples **A** (published) and B (unpublished) in one basket would "spoil both data sets." I cannot find any argument supporting this conclusion. For me the opposite conclusion is evident: "Throwing both data sets in one basket" repaired Gauquelin's past faults. In my reply to de Jager's letter I asked for the missing logical premises of his conclusion, but I received no answer to that question. Providing missing premises needed for understanding conclusions (de Jager) may be more difficult than providing missing data (Gauquelin) needed for empirical evidence.