

## Comments on Serge Daigno's two astrological studies

The author should be congratulated on the very clear setting out of his results. Such clarity is sadly lacking in much published astrological work. As a former professional astrologer I was encouraged by the suggestion that his results might give unbelievers a headache.

### Author's results

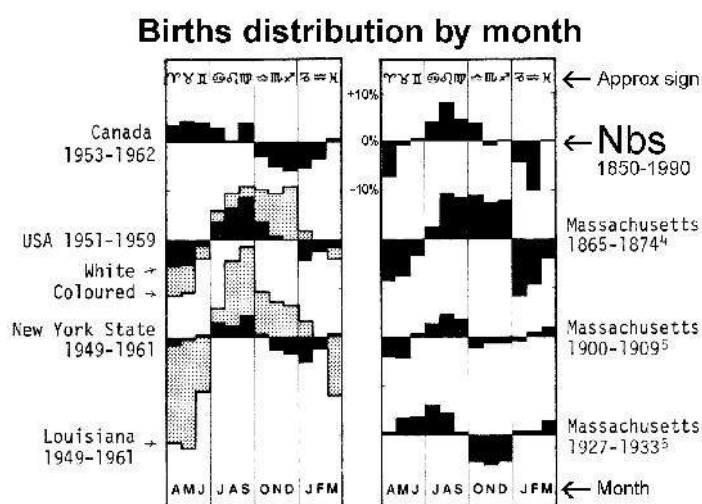
*Study 1.* 1684 chess grandmasters born during 1880-1999 were compared with a control group of births on 587,141 dates picked at random from the same period. Conjunctions between Mercury and Venus were significantly more numerous (5.46% for grandmasters vs 3.52% for the controls, calculated from data in Annex 1.1 for 2° orb).

*Study 2 combined results.* 46,550 team players born between 1850 and 1990 who had played for 0+ years in one of four major US team sports (15,026 Major League Baseball, 3691 National Basketball Association, 23,193 National Football League, 4640 National Hockey League) were compared with a control group of births on 690,028 dates picked at random from the same period. Conjunctions between Venus and Mars were more numerous (2.25% for team players vs 1.79% for the controls, calculated from data in Annex 1.3 and from data provided by email). This difference is said to increase with years of practice (conclusions page), but the frequencies in Annex 1.3 confirm this only for 2+ (2.36%), not for 6+ (1.76%).

### Comments

A problem that has been stressed in the literature for many years, for example by Paul Meehl in *Psychological Reports* 66, 195-244, 1990, is that huge sample sizes will inflate the most trivial differences to impressive statistical significance. For example Gauquelin used large samples, often with highly significant results such as  $p < 0.0001$ . But statistical significance is not the same as practical significance. If you made one hundred predictions based on G's best results, on average only four of them would be better than tossing a coin — and even then you should (on G's figures) forget the 99.994% of the population to whom his results did not apply in the first place. Their high significance was totally useless in the consulting room.

Huge sample sizes are so sensitive to trivial differences that extreme care must be taken to match the demographics of the control group to that of the experimental group, otherwise it will be difficult to tell whether observed differences are due to astrology or demographic mismatches. This is a problem because the observed monthly distribution of births varies between countries, between periods, and even between different ethnic groups in the same country.



Plot shows some published mostly US distributions versus the control (Nbs) distribution, the last after conversion to (given percent – mean percent of 8.333), expressed as a percentage of 8.333, so a given percent of 8.333% is plotted as 0% and 9.17% is plotted as +10%. Here 8.333% is of course 100%/12 months. Each distribution has been corrected for the differing number of days per month, and each is plotted to the same scale. The nearest tropical sign equivalents are at the top.

The differences per month for different countries, states, periods and ethnic groups are clearly evident. There is a broad similarity with the Nbs distribution, but under such sensitive conditions any mismatches will be magnified, ready to be (wrongly) interpreted as astrological. The control distribution needs to be based pro rata on the various birth places and periods.

The problem is made worse because the occurrence of conjunctions is non-uniform due to planetary retrograde motion, which makes the accurate matching of controls especially important, and because the demographic data needed to accurately match a given experimental group is almost never available. This means that observed small differences can almost never be interpreted. One answer of course is differences too large to be misinterpreted!

An associated pitfall that Gauquelin occasionally (and inadvertently) did not avoid was to calculate chi-squared values using a control group for the expected values instead of the theoretical values. The control group will be subject to sampling error that does not apply to theoretical values, which the chi-squared test does not allow for, so the apparent significance level is made artificially higher than it should be.

One standard way to overcome this problem, and also the problem of mistaking statistical significance for practical significance, is to look at the actual association between results in a 2 x 2 contingency table, which for grandmasters would look like this (data from Annex 1.2):

		Grandmasters		Nbs		
Mercury-Venus conjunction? orb 2°	Yes	a	92	b	20691	
	No	c	1592	d	566450	
			1684		587141	= 0.006

and like this for 0+ team players (data from Annex 1.3 and from data provided by email):

		Team players		Nbs		
Venus-Mars conjunction? orb 2°	Yes	a	1046	b	12355	
	No	c	45504	d	677673	
			46550		690028	= 0.008

From such a table the measure of association (phi) can be calculated by the equation given in any statistics book, namely  $\phi = (ad-bc) / \sqrt{((a+b)(c+d)(a+c)(b+d))}$ . Values of phi can range from 1 (indicating perfect association as in inches vs centimetres), through 0 (indicating zero association) down to -1 (indicating perfect inverse association as in day vs night). In this case the value of phi provides an immediate answer to the key question — *to what extent* is being a chess grandmaster or team player associated with a given astrological conjunction?

By way of an example, the association between answers to a single question in an IQ test and the resulting IQ score is typically around  $\phi = 0.15$ , which means that the answers to a single question will say little about IQ. But when many questions are combined together the result becomes much more meaningful (eg  $\phi > 0.8$  for >50 questions), and is limited only by how many suitable questions the average person can cope with, and by the funding needed to derive suitable questions in the first place. So a small association is not necessarily bad news.

For grandmasters my statistical software calculated these results from the data in Annex 1.2:

Planet	Venus conjunct planet orb 2°	P	Same with orb 6°	P
Mercury	0.006	.00002	0.003	.034
Sun	0.000	.88	0.001	.55
Mars	0.003	.04	0.001	.36
Jupiter	0.001	.42	-.002	.18
Saturn	-.000	.96	-.000	.83

And these results for team players combined, calculated from the data in Annex 1.3 and the sample size data provided by email, all with orb 2°

Planet	Venus conjunct planet 0+	P	Same for 2+	P	Same for 6+	P
Mercury	0.001	.44	0.001	.56	-.011	.00000
Sun	0.002	.11	0.001	.53	-.010	.00000
Mars	0.008	.00000	0.008	.00000	-.000	.74
Jupiter	-.002	.12	-.001	.24	-.007	.00000
Saturn	0.002	.13	0.002	.12	-.005	.00002

(for comparison: Mars effect  $\phi = 0.04$ , sun sign self-attribution  $\phi = 0.08$ )

All values of  $r$  are close to zero, indicating negligible association between variables. The effective values of  $p$  have of course to be corrected for the number of results, but even then they are most likely meaningless due to inflation by a mismatched control group.

Some clues may be provided here by the symbolism. As the author points out, the traditional meaning of Mercury conjunct Venus (*mentally harmonious, charming disposition*) is difficult to reconcile with the mental aggression and challenge presumably required in chess.

Similarly the traditional meaning of Venus conjunct Mars (*sensitive, easily angered, efficiency is reduced by over-active emotions*) is difficult to reconcile with the robustness and aggressive energy presumably required in a team sport.

Furthermore, as they stand, the  $r$  values for team players indicate that Venus conjunctions are generally favourable for 0+ and 2+ team players ( $r$  is generally positive), but not for 6+ team players ( $r$  is negative). This change in direction is difficult to explain by astrology, but is plausibly explained by the reduced sample size (6+ players with a Venus conjunction are typically half as numerous as 2+ players), which will increase the sampling error and most likely shift the years sampled, thus creating an increasing mismatch with controls generated by a random sampling from 1850-1990 despite adjustments to match the Nbs distribution

The same points apply to the results of each individual team sport.

One final problem related to practical application is the low base rate of conjunctions. Even if the claimed conjunctions orb  $2^\circ$  were very highly associated with being a grandmaster or team player, only about 5% of grandmasters and 2% of team players would actually have one. The remaining 95% and 98% would seem to be left out in the cold. Indeed, according to astrology they could now be denied being a grandmaster or team player, while ordinary people who happened to have one could claim automatic status. All of which might be problematic.

### **Conclusion**

The above considerations weigh against an interpretation of the results in terms of astrology.

The apparent connections with astrological conjunctions are most likely an artifact of low base rates and an insufficiently matched control group.

Even if the connections were shown by replication to be genuine, the level of association is too low to have any practical utility, although it might be of scientific interest.

One problem shared with current research is the focus on statistical significance instead of practical significance. As pointed out by Meehl, who at the time was President of the American Psychological Association, if the sample size is large enough (he gives a case of  $N = 57,000$ , considerably smaller than the present controls), the relation between any two variables picked at random will *always* give significant  $p$  values. Sooner or later even the most unlikely-to-be-related variables will be significantly if uselessly related. Astrologers who pay attention only to statistical significance (which is all of them) will immediately see such significance as evidence for astrology when of course it may be nothing of the sort.

After so much hard work such problems need not be a setback because the associated capability is impressive (it has been a privilege to see it at work), and suitable changes to parameters could be worth following up. Which of course is easy to say. Those faced with actually doing the work might disagree.

June 2017