

## Intelligence models for the Dutch Raven Data

***Citation for published version (APA):***

Nienhuys, J. W. (1983). *Intelligence models for the Dutch Raven Data*. (Eindhoven University of Technology : Dept of Mathematics : memorandum; Vol. 8301). Technische Hogeschool Eindhoven.

***Document status and date:***

Published: 01/01/1983

***Document Version:***

Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

***Please check the document version of this publication:***

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

***General rights***

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.tue.nl/taverne](http://www.tue.nl/taverne)

***Take down policy***

If you believe that this document breaches copyright please contact us at:

[openaccess@tue.nl](mailto:openaccess@tue.nl)

providing details and we will investigate your claim.

696974

INTELLIGENCE MODELS FOR  
THE DUTCH RAVEN DATA

by

J . W . NIENHUYS

The author is Scientific Officer  
at the Mathematics Department of  
the Eindhoven University of Technology.  
Address: P.O. Box 513, 5600 MB Eindhoven,  
The Netherlands

## SUMMARY

Zajonc and Marcus developed in 1975 a model to explain the striking relation between birthorder and intelligence shown by the Dutch Raven Data of Belmont and Marolla.

I show that this model suffers from errors in logic, computation and methodology. No meaningful relation exists between the model and the data. I argue that other explanatory variables, such as social grouping, must be considered as well. Moreover, models for the Dutch Raven data are so easy to construct, that any model is questionable.

## 1. INTRODUCTION

Belmont and Marolla <sup>1)</sup> published an investigation into the relation between intellectual performance and family composition. By family composition I mean father's occupation, total number of children and one's own birth rank in the family of origin. Intellectual performance means the test score on a version of the Raven Progressive Matrices test, used by the Netherlands Department of Defense during the pre-induction physical examination <sup>2)</sup>. Throughout this paper I denote the average score of individuals with birth rank  $i$  from a family of size  $j$  by  $M_{ij}$ . Roughly speaking,  $M_{ij}$  decreases with increasing  $i$  and  $j$ .

Zajonc and Markus <sup>3)</sup> expanded a remark of Belmont and Marolla about a possible environmental explanation for the way  $M_{ij}$  depends on  $i$  and  $j$ .

This article won the authors the Socio-Psychological Prize of the AAAS in 1975 <sup>4)</sup>.

The article of Zajonc and Markus contains two parts that I consider here.

Firstly, an explanation of the Belmont-Marolla data and the presentation of an environmental model, the so-called Confluence Model, to account for some details of the Belmont-Marolla data.

Secondly, a computation to show that the model provides a satisfactory fit to these data.

In the sequel, I summarize these two parts and analyze the computation. I shall show that the computations do not mean very much, and that the seemingly nice fit hides large discrepancies between model and reality.

In the remainder I shall examine the Belmont-Marolla data, and I shall argue that "father's occupation" is a variable that cannot be missed in feasible models. Finally, I will show that it is very easy to propose other theories that account equally well for the Belmont-Marolla data.

## 2. THE CONFLUENCE MODEL

Zajonc and Marcus suggest that the Belmont-Marolla data exhibit the following two phenomena, aside from the ones mentioned already:

Firstly, last-born children, including only-borns score less well than one might expect by extrapolation from the rest of the population. I will call this the Benjamin-effect <sup>5)</sup>.

Secondly,  $M_{ij}$  is a quadratic function of  $i$  <sup>6)</sup>.

The Confluence Model consists of three hypotheses.

(A) The absolute intelligence <sup>7)</sup> of humans develops according to the following differential equation:

$$\frac{d}{dt} I(t) = \alpha(t) \frac{d}{dt} S(t)$$

$$I(0) = S(0) = 0$$

where  $I(t)$  denotes absolute intelligence at age  $t$ ,  $\alpha(t)$  denotes the intellectual environment experienced at age  $t$ , and  $S$  is a fixed function.

The graph of  $S$  resembles vaguely a logistic growth curve <sup>8)</sup>.

(B) The function  $\alpha$  in hypothesis (A) is the average of the absolute intelligences of all members of the family, parents included, at the time the individual has reached age  $t$ .

For example, in a family with two adult parents of age 30, and two children aged 7 and 4, the intellectual environment  $\alpha$  would equal

$$\frac{1}{4} ( I(30) + I(30) + I_1(7) + I_2(4) )$$

where  $I_1$  and  $I_2$  denote the absolute intelligence functions of the oldest and the youngest child.

(C) The combination of (A) and (B) implies that prolonged social intercourse with younger, hence absolutely less intelligent children has a depressing effect on one's own intellectual development. Zajonc and Marcus state<sup>9)</sup> that this is also suggested by casual observations. However, the absence of this depressing effect is even more depressing. Indeed, last-borns have no younger brothers or sisters and according to hypothesis (C), this results in the lack of the edifying experience of teaching younger siblings. Hypothesis (C) accounts for the Benjamin-effect.

The authors illustrate (A) and (B) with a few computations with fictitious numbers. These computations only show that in nine-child families,  $\alpha$  is quadratic as function of the number of siblings born, and that this function is different if differently sized gaps between births are assumed.

Nothing about a quadratic trend of the cumulative effects at age 19, as function of birth order, is shown.

The authors also show that their model makes predictions that conform to the Belmont-Marolla data. This is done as follows.

A second child in any family passes through different stages. First it is a second of two, immediately after birth. When a next sibling is born, it becomes a second of three, then a second of four and so on.

The authors assume that there exists a second-of-two environment,  $\alpha_{22}$ . From hypothesis (A) and  $M_{22}$ ,  $\alpha_{22}$  can be computed. This number represents equally well the average environment of a second-of-three during its first four years as the average environment of a second-of-two. Now again by (A) and  $M_{23}$ ,  $\alpha_{23}$  can be computed, and so on. Environments that should be vastly different according to (B) are identified over and over again. Moreover, in this computation it is apparently assumed that having no younger brothers or sisters at the age of zero to two results in a handicap, by (C). In this way almost everybody in the whole world has had such a handicap.

Summarizing, hypotheses (A), (B) and (C) are violated in every step of the

computations. Also, the more stages a hypothetical child has to pass, the more errors will pile up in the computation of its last environment. So the environments  $\alpha_{ij}$  obtained in this way are useless for checking the validity of the Confluence Model.

Hypothesis (B) says that all individuals in a given family will experience the same environment. The computations just sketched yield as many different environments  $\alpha$  as there are members of the family (parents excluded). These should all be equal, or at least all but the last one, because of the Benjamin-effect.

Even though the numbers  $\alpha$  are computed by violating hypothesis (B), one might check whether they do or do not depend on birth rank. This is what Zajonc and Marcus do.

They begin to observe that the Belmont-Marolla data can be approximated with a formula of the form

$$a_0 + a_1i + a_2j + a_3i^2 + a_4\lambda$$

where  $\lambda = 0$  if  $i=j$  and 1 otherwise, and where  $a_0, a_1, a_2, a_3, a_4$  represent suitably chosen numbers. Such an approximation yields a surprisingly nice fit: only 3% of the variance is unaccounted for.

For later reference I call this computation X.

Subsequently the authors observe that computation X applied to their  $\alpha$ 's leaves 31% of the variance unaccounted for, and that the coefficients  $a_1$  and  $a_3$  do not differ significantly from zero in this case. They comment that "family size ... necessarily depresses intellectual environment".

Altogether this does not support the Confluence Model very well. Even without logical flaws in the computation of the  $\alpha$ 's it is not clear why computation X should be used at all for the  $\alpha$ 's, rather than approximations with other functions of  $i$  and  $j$ . Moreover a tenfold increase in unexplained variance should not be ignored.

I checked the computations and I found that they are wrong. In fact, 72% of the variance is unaccounted for. Closer investigation yields that the first of a family of nine lives in an environment that is about 14 standard deviations below normal, at least, according to this computation. Deletion of the offending environment and multiple regression on the remaining 44 numbers give the figures that Zajonc and Markus disclose about their computation.

Actually in this way results can be improved further, for instance by deletion of all  $\alpha_{ij}$  with  $j-i$  more than 5. In that case only 7% of the variance is unaccounted for.

Summarizing, I find that logical and methodological and other errors in the computations destroy any support that the Belmont-Marolla data give the Confluence Model.

Far from being supported by 386,114 data points, the Confluence Model is not quantitatively supported by any measurement mentioned by Zajonc and Markus.

### 3. A CLOSE LOOK AT THE BELMONT-MAROLLA DATA

What parameters do describe the Dutch Raven data of Belmont and Marolla adequately? Does the only-born child have a handicap of its own or is it merely an instance of a last-born child? Henceforth I call the hypothetical handicap of an only-born the Isaac-effect.

I try to answer these questions by fitting simple functions of  $i$  and  $j$  to the data.

Dr. Belmont has kindly provided me with more detailed data, namely for every family composition (see introduction) the size of the corresponding subpopulation and its distribution over class scores 1 through 6.

These subpopulations differ in size by a factor 20, so weighted approximations are called for.

Repeating computation X, but with a term for the Isaac-effect, with weighted approximations and excluding familysize 9+, I found 98% of the variance accounted for. This seems very good, but it is very poor. If the model were correct, the sum of the squared residues should follow a chi-squared distribution with 30 degrees of freedom (all dependent variables are standardised at variance 1). The squared residues sum equals 96.3 and this gives a tail probability of  $7 \times 10^{-9}$ , in other words this model must be rejected. Similar computations with the three professional groups raised a suspicion that the three groups actually follow different distributions. To test this, I considered the following hypothesis  $H_0$ :

Raven Ability RA is a variable that is normally distributed in each subpopulation consisting of men with the same family composition. Each subpopulation has its own mean and variance. The men tested by the Dutch Army in the years 1963-1966 form a simple random sample from these subpopulations. The Raven Test assigns each man a class score, depending on his Raven Ability, as follows:

RA less than  $r_5$  is assigned class score 6, RA not less than  $r_1$  is assigned class score 1 and the remaining class boundaries are  $r_2$ ,  $r_3$ , and  $r_4$ .

Moreover I considered similar hypotheses  $H_1$ ,  $H_2$  and  $H_3$ , but these refer to the non-manual, manual and farm populations separately.

Observe that the existence of the class boundaries forms part of the hypotheses, but that this does not hold for the values of these boundaries. This makes the test of the hypothesis less sensitive to deviations from normality.

To test these hypotheses, I tried to choose for each subpopulation a mean and a variance and moreover the three ratios

$$r_4 - r_5 : r_3 - r_4 : r_2 - r_3 : r_1 - r_2$$

in such a way that the differences between the actual sample distributions and the theoretical distributions became as small as possible.

The difference between actual and theoretical distribution I measured by  $\chi^2$ . This amounts in case of hypothesis  $H_0$  to minimizing a function of 219 variables.

The tail probabilities for hypotheses  $H_0$ ,  $H_1$ ,  $H_2$  and  $H_3$  are <sup>10)</sup>  $3 \times 10^{-8}$ , 0.95, 0.24 and 0.07 respectively. It is not difficult to argue that the last of these numbers should read 0.14 or maybe even more <sup>11)</sup>.

Consequently I reject  $H_0$ , but I see no compelling reason to reject  $H_1$ ,  $H_2$  and  $H_3$ .

This indicates that the different populations should not be put together: they follow different distributions.

For example, Raven Class 6 seems too large in the non-manual population or too small in the two other populations.

The non-manual population shows another remarkable phenomenon: among first-borns (and to a lesser degree also among last-borns) the variances of the Raven Ability (as computed in the course of testing  $H_1$ ) increase regularly with increasing family size. For first-borns the range is from 0.93 (family size 1) to 1.23 (family size 8). For other birth orders and other professional groups there are no such striking regularities.

What parameters describe the Belmont-Marolla data?

To answer this question, I did computation X again for the three professional groups separately. I used the means obtained from the calculations for testing  $H_1$ ,  $H_2$  and  $H_3$ . Moreover I used additional independent variables, namely  $ij$  and  $j^2$ , but I threw out variables that did not reach the 0.05 level of significance. Of course this is somewhat arbitrary, and one might get different results if one takes  $(i+j)$  and  $(i-j)$  as independent variables, instead of  $i$  and  $j$ . Especially after throwing away insignificant terms one might then arrive at a different answer. At this point one suffers from the

lack of a theoretical foundation, that can tell us what to expect.

Anyway, after this computation I find:

- (i) The Isaac-effect and the Benjamin-effect are significant at the 0.001 level in all three populations, except the Benjamin-effect in the farm group.
- (ii) There <sup>are</sup> only significant birth rank effects ( $i$  and  $i^2$ ) in the farm group and I found no significant family size effects in the farm group.
- (iii) The way the approximations depend on  $i$ ,  $i^2$ ,  $j$ ,  $j^2$ ,  $ij$  in the manual group and in the non-manual group is quite complicated and also different in both groups. This is sketched in Figure 1. The curved lines connect points where the sum of the contributions of the constant term and the terms with  $i$ ,  $i^2$ ,  $ij$ ,  $j$ ,  $j^2$  are equal: the so-called level curves.

Figure 1 is again evidence that social grouping must be considered in explanations of family size and birth rank effects in results of intelligence tests.

(iv) The tail probabilities for the three populations (the probabilities that the observed sums of squared residuals or <sup>larger</sup> ones, occur if one assumes that the model is correct) are 2% (non-manual), 3% (manual) and 17% (farm). This is not much (though better than  $7 \times 10^{-9}$ ), but not surprising either. There is, to my knowledge, no theory that predicts the functions, or the magnitude, or even the sign of the coefficients to be used.

Belmont, Stein and Susser <sup>12)</sup> reported height in a part of the population.

Doing again computation X with height as one more extra variable, I found:

- (v) Among family sizes 1 through 6, the Isaac-effect and the pure family size effects ( $j$  and  $j^2$ ) become insignificant in the whole population if height is taken as independent variable. The tail probability is 28%. This is evidence for the proposition that the Isaac-effect is different from the Benjamin-effect.

## 4. SOME FANCY EXPLANATIONS

I show now how easy it is to propose theories that make vague predictions about the way the numbers  $M_{ij}$  should depend on  $i$  or  $j$  or both.

To prove such a theory one would have to make a retrospective prediction that fit the Belmont-Marolla data better than the trial-and-error approximations of the last section. I trust the reader will appreciate the difficulty of refuting or proving any of these theories.

## THEORY I (Who-likes-the-army theory)

Young people do not like the army. The prejudice a boy has against the army is directly proportional to the number of brothers already serving or having served in the army. The achievement on the Raven test during the pre-induction examination depends linearly on the amount of prejudice against the army.

This theory seeks the explanation in the test situation and predicts a linear relation between  $M_{ij}$  and  $i$ , maybe flattening out after  $i = 6$  because of the exemption for those who had three brothers serving already.

## THEORY II (Playground theory)

Around farms and to a lesser degree in houses of white collar workers there is relatively much space. The amount of space in and around the house available as playground is an important environmental variable with respect to exploratory behavior and the development of intelligence.

This theory explains why  $M_{ij}$  decreases with increasing  $i$ , and also why - aside from the Benjamin-effect- farm families and small non-manual families show no family size effects.

This theory seeks an explanation in a well-definable environmental variable with the character of a common resource, that acts a young age.

## THEORY III (Large-families-have-old-parents theory)

Children in the sample with high birth ranks have older parents than children with low birth ranks and same family size. As educational opportunities improve continuously, parents of high birth ranks have received less formal education, on the average. This theory predicts that  $M_{ij}$  decreases with increasing  $i$ . It seeks an explanation in the fact that the different subpopulations cannot be considered as comparable simple random samples.

## THEORY IV (Economic theory)

Money is an environmental variable with the character of a common resource. Children with high birth ranks or from large families get less of this resource. Pursuing formal education after the maximum age for compulsory education (about 15) costs money even if it is only because someone that sits in school cannot contribute to the family income.

Van Heek <sup>13)</sup> found that the decision to let a child go to any particular type of school is not affected very much by family size or birth rank. But what happens at the ages of 12 and 15 is not necessarily closely related.

Like theory II, this theory is based on the idea of a common resource, but this resource acts between the age of 15 and 18. It assumes that a few years of formal education has some effect on the Raven Ability as measured by the army.

## THEORY V (Mixture theory)

There are 6.5 times as many first-borns from non-manual families of two, compared to farm families. For families of eight this ratio is 1.4. This means, even if in neither social group  $M_{ij}$  was dependent on  $j$ ,  $M_{12}$  and  $M_{18}$  for the whole population would differ, because the farm group scores lower than the non-manual group.

For other ways of dividing the population into two or more groups with

different family composition patterns and different average scores, such observations might also explain a lot. But were to look for such subdivisions? (Catholic-Remainder? Rural-City? Skilled-Unskilled? )

I do not believe in any of these explanations, nor in any other, including easy-to-make unpleasant sounding genetic explanations. These theories merely serve to demonstrate that making models is so easy, that no model deserves consideration until there is some hard proof for it.<sup>14) 15)</sup>

## NOTES AND REFERENCES

- 1) Belmont, L., and F.A. Marolla, Birth Order, Family Size and Intelligence, Science 182 (1973), p. 1096-1101.
- 2) All men born in 1944 through 1947 that were not exempt or in an institution were tested. Each person was assigned a class score as follows: (number of correct answers, followed by score in brackets):  
 0 - 9 (6); 10 - 16 (5); 17 - 23 (4); 21 - 24 (3); 25 - 29 (2); 30 - 40 (1).  
 The differences between averages of groups with adjacent family sizes and birth orders are in the order of 0.1 score point, which is small compared to the standard error of measurement. This means that the discussion is rather theoretical: the differences under discussion are meaningless on the level of individuals but the test is designed and used to measure gross differences between individuals. This makes differences between populations hard to interpret.
- 3) Zajonc, R.B., and G.B. Markus, Birth Order and Intellectual Development, Psych. Rev. 82 (1975), p. 74 - 88.
- 4) Zajonc, R.B., Family Configuration and Intelligence, Science 192 (1976), p. 227 - 236.
- 5) The names Benjamin and later on Isaac are taken from a well-known literary work where these are the names of a youngest son and an only born son respectively.
- 6) By quadratic is meant that the second derivative does not change sign.
- 7) Zajonc and Markus do not explain what absolute intelligence is, how it could be measured and on what basis the growth of this hypothetical quantity is described by the function they give.

8) Solutions of the differential equation of logistic growth and of the equation given by Zajonc and Markus have in common that they tend to a limit that is the same for all solutions, and in case of the equation of Zajonc and Markus, also the same for all functions  $\alpha$  that have a given limit. Such equations cannot describe the basic idea of the Confluence Model, which assumes that the limit depends on the whole life history. Another feature of these equations is that the speed of growth also depends on the size of whatever is growing. From the computations Zajonc and Markus make can be seen that they compute the rate of growth without using the size reached. Maybe these two errors cancel?

9) loc. cit. p. 84.

10) The sum of all 108  $\chi^2$  does not depend linearly or otherwise nicely on the parameters to be chosen. Nevertheless I counted every parameter chosen as a loss of one degree of freedom. So the sum is assumed to follow a  $\chi^2$  distribution with 321 degrees of freedom. It is 472.8.

11) The farm variances differ on the average less than 4% from their central value. So the part of hypothesis  $H_3$  that states that each subpopulation has its own variance is too strong. A hypothesis  $H_3'$ , where all variances are assumed equal, gives a tail probability of .14.

Moreover, in the farm case, the subpopulation corresponding to  $M_{88}$  ( $N=351$ ) makes a large contribution to the sum of the  $\chi^2$ . This subpopulation is strongly bimodally distributed over the class scores. It is also the only bimodally distributed subpopulation among all professional groups, family sizes and birth orders less than 9. One almost suspects a clerical error.

12) Belmont, L., Z.A. Stein, M.W. Susser, Comparison of associations of birth order with intelligence test score and height, *Nature* 255 (1975), p. 54 - 56.

I have used the graph in that article. The corresponding numbers were not available.

13) Heek, F. van, et al., Het Verborgen Talent, Meppel, Boom, 1968.

14) Program texts in Burroughs Extended Algol, intermediate results of computations and about 14 pages with summaries of the results of the computations can be obtained from me.

15) I thank L. Belmont for detailed numerical data on the test scores, P.W.J. Groen of the Netherlands Department of Defense for the information in the first part of note 2), and P.A. Vroon for moral support.

## LEGEND FOR FIGURE 1

Level curve diagram for Raven performances of the Non-Manual Group (left) and the Manual Group (right). The best approximations, rather than the actual data are plotted here. Benjamin and Isaac-effects have been discarded. The + marks indicate the points corresponding to the actual data. The values on any curve may be thought of as multiples of the population's standard deviation. Hence one might say that two adjacent lines differ by about  $1/3$  IQ-point.

Although it is not clear how to interpret such small differences between populations, these graphs suggest that social grouping must enter into explanations of birth rank effects on intelligence.

