Psychology Science, Volume 45, 2003 (2), p. 254-262

Testing the statistical association between family names and success in certain athletic disciplines in men called 'Smith' or 'Tailor'

MARK STEMMLER[1], GÜNTHER BÄUMLER[2]

Summary

In respect to the medieval practice of giving surnames based on ones profession and based on the hereditary transmission from generation to generation, Bäumler (1980) suggested a genetic-social theory of assortative distribution of traits of body build such as height, weight, and stature in a population of men called 'Smith' (German: Schmied) and 'Tailor' (German: Schneider). From this theory the hypothesis was deduced that among the top ranking athletes of the 'heavy weight' branches of athletics, which require body strength and body height, there are relatively more persons that go by the name of Schmied than in the 'light weight' branches of athletics, which are more stamina demanding, where more persons go by the name of Schneider. At the same time, complementary hypotheses assuming less Schneiders among the heavy weight branches than expected and less Schmieds among the light branches of athletics than expected under the null hypothesis of no statistical association were also tested. All hypotheses were empirically supported by applying the data of two independent samples.

Key words: Analysis of contingency tables, configural frequency analysis (CFA), chi-square decomposition, biprediction, Kimball test, genetic-social theory, heredity, surnames, track-and-field athletics, body build

[1] PD Mark Stemmler, Ph.D., University of Erlangen-Nuremberg, Department of Psychology, Bismarckstr. 1, D-91054 Erlangen, E-Mail: mkstemml@phi.uni-erlangen.de

[2] Prof. Dr. Günther Bäumler, Technical University Munich, Department of Sport Science, Connollystr. 32, D-80809 Munich, http://www.sport.tu-muenchen.de

1. Introduction

This article is a sequel of a paper published by Bäumler (1980). Whereas the first paper was dealing with genetic/biological-social differences in physique (e.g., height, weight, solidity) in men called 'Smith' and 'Tailor' the following analyses are based on a different study performed by Bäumler (1984) and are investigating the association between the above mentioned family names and success in certain branches of athletics. First, a sample of N = 6099 subjects was selected and stratified according to their last names: (1) individuals with the last name of 'Smith' (in German 'Schmied', including all derivations like 'Schmidt' etc.), (2) individuals with the last name of 'Tailor' (in German 'Schneider'), and (3) individuals with other last names. These individuals were taken from a listing of the best athletes of the year 1977 in Germany belonging to ten different track-and-field associations and encompassing 16 different track-and-field disciplines. For the data analyses the track-and-field disciplines were classified into three different athletic categories: 'heavy-stature' (including athletes like shot-putters, discus throwers, etc.), 'medium-stature' (including athletes like pole-vaulters, 100 meters runners, etc.) and 'light-stature' (including athletes like 1500 meters runners, 10000 meters runners, etc.). The frequencies for names like 'Smith' and 'Tailor' were taken from this list of athletes.

The data (cf. Bäumler, 1984) revealed that among the successful athletes in the heavy-stature disciplines 1.7 percent were 'Smiths', in comparison to only 0.09 percent which were called 'Tailors'. The same picture evolved in the medium-stature disciplines, here 1.23 percent were called 'Smith', whereas only 0.22 percent were called 'Tailors'. Another picture evolved in the light-stature disciplines: Here, 0.98 percent were called 'Smiths' and 1.14 percent were 'Tailors'. This finding, which was statistically significant, was discussed before (Bäumler, 1984; Bäumler and Lienert, 1987) in terms of the above mentioned 'genetic-social theory' which is based on a fact which has its origin in the 17th and 18th century where firstly, the use of last names became legally required and secondly, where the choice of last names was often associated with the profession and the body build of the person in question. Thus, the persons called 'Smiths' orginally stem from families of blacksmiths, a profession requiring a certain physical aptitude to cope with the hard physical labour. The same was true for the "light-weight" 'Tailors' who stem from families of tailors who were, like the blacksmiths organized in their own guilds; all new apprentices were recruited through these guilds, a fact which kept the business in the family. Moreover, the social life in former times was also determined by the guilds, including various festivities of the guilds, increasing the chance that the daughter of a smith will marry another smith and so forth. The above described procedures resulted in an 'inbreeding of populations' with a strong dominance of a certain family name connected to a certain body build. Therefore, by applying this genetic-social theory of a combined heredity of body build and family names it can be hypothesized that family names and physique were handed down from generation to generation up to the present (it is estimated that there are about 8 to 20 generations between the original ancestors and the present bearers of these names; Bäumler, 1980) and that these different physiques influence the success in different branches of athletics (cf. Bäumler, 1984; Bäumler and Lienert, 1987). In this paper also the complementary hypotheses, which assume less Schneiders among the heavy weight branches than expected and less Schmieds among the light branches of athletics than expected under the null hypothesis of no statistical association will be tested.

2. Testing the genetic-social theory globally

To test our hypothesis inferentially a 3 by 2 table was constructed. For reasons of parsimony the heavy and medium stature disciplines were collapsed (see Table 1).

Table 1:
Frequencies of successful athletes according to their last names 'Smith', 'Tailor' and all other names by two categories of track-and-field (heavy-and-medium and light)

	Physical Statures				
	heavy-and-medium		light		
Successful Athletes	f(o)	f(e)	f(o)	f(e)	Column marginals: ∑
Smiths	a = 70	65.49	b = 12	16.51	A = 82
Tailors	c = 8	17.57	d = 14	4.43	B = 22
Other last names	4793	4787.94	1202	1207.06	5995
Row marginals: ∑	C = 4871		D = 1228		N = 6099

Note. Global $\chi^2 = 27.5$, df = 2, $p < .05$. The letters refer to Kimball's formula. f(o) = observed frequencies; f(e) = expected frequencies.

The columns represent the categories of heavy-and-medium-stature and light-stature disciplines. The rows are divided into athletes with the names of 'Smith', 'Tailor', and other last names. We expected that 'Tailors' are most frequently successful in the light-stature disciplines, whereas the 'Smiths' can be found preferably in the heavy-and-medium-stature disciplines. A configural frequency analysis (CFA) was applied (Lienert and Krauth, 1975; von Eye, Spiel, and Wood, 1996; von Eye, 2002) which compares observed with expected frequencies. The expected frequencies were calculated on the assumption of independence between the rows and columns. All analyses were done with the program SICFA (Lautsch and von Weber, 1995). The global chi-square was 27.5 with df = 2 $p < .05$, indicating a significant association between the rows and columns:

$$\chi^2 = \sum_{i=1}^{k} \frac{(f(o) - f(e))^2}{f(e)} =$$

$$= \frac{(4793 - 4787.94)^2}{4787.94} + \frac{(1202 - 1207.06)^2}{1207.06} +$$

$$+ \frac{(70 - 65.49)^2}{65.49} + \frac{(12 - 16.51)^2}{16.51} + \frac{(8 - 17.57)^2}{17.57} + \frac{(14 - 4.43)^2}{4.43} +$$

$$+ \frac{(4793 - 4787.94)^2}{4787.94} + \frac{(1202 - 1207.06)^2}{1207.06} =$$

= 27.5 with df = 2, $p < 0.05$.

In a next step, two assumptions were tested simultaneously (Lienert and Netter, 1987): First, that 'Smiths' can be found predominantly in the heavy-and-medium-stature disciplines and secondly, that 'Tailors' can be found in the light-stature disciplines. This was obtained by applying an orthogonal chi-decomposition according to Kimball (1954; cf. Stemmler, 1994; cf. Müller and Stemmler, 2002). The respective null hypothesis states that 'Smiths' and 'Tailors' are equally distributed among the physical statures. The formula of Kimball for our 3 by 2 table looks as follows:

$$\chi^2 = \frac{[A(Cd - Dc) - B(Cb - Da)]^2}{\frac{ABCD(A+B)(C+D)}{N}} =$$

$$= \frac{[82(4871\times 14 - 1228\times 8) - 22(4871\times 12 - 1228\times 70)]^2}{\frac{82\times 22\times 4871\times 1228\,(82+22)\,(4871+1228)}{6099}} =$$

$$= 5.09 \quad \text{with df} = 1;\ p < 0.05$$

This significant chi-square shows that different proportions of 'Smiths' and 'Tailors' can be found in the two athletic categories. Therefore, the underlying hypothesis of a different success rate of 'Smiths' and 'Tailors' based on different track-and-field disciplines was supported. This hypothesis was supported although all the 'Non-Smiths' and 'Non-Tailors' were taken into consideration. Another question is whether the 'Smiths' and 'Tailors' equally contribute to this finding or whether the association between success in certain athletic disciplines and last name is the same for the two 'professions' under investigation. For the investigation of this issue the following analyses were performed.

3. A differential test of the genetic-social theory

Table 1 was transformed such that only a fourfold table evolved in which the proportions of 'Tailors' and 'Other last names' with regard to the two stature categories was tested inferentially. Table 2a presents the transformed cell frequencies.

Table 2a:
Comparing 'Tailors' and athletes with other last names with regard to their physical statures.

	Physical Statures				
	heavy-and-medium		light		
Successful athletes	f(o)	f(e)	f(o)	f(e)	Column marginals: ∑
Tailors	8	17.6 A	14	4.4 T	22
Other last names	4863	4853.4	1214	1223.6	6077
Row marginals: ∑	4871		1228		N = 6099

Note. Global $\chi^2 = 26.0$, df = 1, $p < .05$. f(o) = observed frequencies; f(e) = expected frequencies. A = antitype; T = type.

By comparing the observed with the expected frequencies a total chi-square of 26.0 with df = 1 and p < .05 was obtained:

$$\chi^2 = \sum_{i=1}^{k} \frac{(f(o) - f(e))^2}{f(e)} =$$

$$= \frac{(8-17.6)^2}{17.6} + \frac{(14-4.4)^2}{4.4} + \frac{(4863-4853.4)^2}{4853.4} + \frac{(1214-1223.6)^2}{1223.6} =$$

$$= 26.0 \quad \text{with df} = 1;\ p < 0.05.$$

One type and one antitype (von Eye, 1990, 2002) were detected. Types represent cells with significantly more observations than expected under the null hypothesis, and antitypes represent cells with significantly less observations than expected. The identified type revealed that more 'Tailors' than expected can be found in the light-stature category; the identified antitype suggested that there are less 'Tailors' than expected in the heavy-and-medium-stature category.

Table 2b:
Comparing 'Smiths' and athletes with other last names with regard to their physical statures.

	Physical Statures				
	heavy-and-medium		light		
Successful athletes	f(o)	f(e)	f(o)	f(e)	Column marginals: ∑
Smiths	70	65.5	12	16.5	82
Other last names	4801	4805.5	1216	1211.5	6017
Row marginals:∑	4871		1228		N = 6099

Note. Global χ^2 = 1.6, df = 1, n.s. f(o) = observed frequencies; f(e) = expected frequencies.

Table 2b lists the corresponding frequencies for 'Smiths'. Unfortunately, the global chi-square misses the level of significance:

$$\chi^2 = \sum_{i=1}^{k} \frac{\left(f(o) - f(e)\right)^2}{f(e)} =$$

$$= \frac{(70-65.5)^2}{65.5} + \frac{(12-16.5)^2}{16.5} + \frac{(4801-4805.5)^2}{4805.5} + \frac{(1216-1211.5)^2}{1211.5} =$$

$$= 1.6 \text{ with df} = 1,\ p = \text{n.s.}$$

However, the cell frequencies are in the expected directions: There are more successful 'Smiths' than expected in the heavy-and-medium-stature category and less than expected in the light-stature category. The analyses of Table 2a and 2b show that today the association between last name and success in certain athletic disciplines is much stronger for 'Tailors' than for 'Smiths'.

4. Validation of the findings in a different sample

In order to validate the above mentioned findings Bäumler, Forstner, and Stemmler (2002) obtained a new sample of successful athletes and their respective surnames. The list consisted of the best athletes of the year 1983 in West-Germany belonging to 15 different track-and-field associations and encompassing 20 different track-and-field disciplines (heavy-and-medium group: shotput, discus throw, hammer throw, javelin throw, decathlon, 100 meters, 200 meters, 400 meters, long jump, 110 meters hurdles, 400 meters hurdles, pole-vault, walking; light group: 800 meters, 1500 meters, 3000 meters, 5000 meters, 10.000 meters, 25.000 meters, marathon). The above mentioned ideas about handing down a special physique in men from generation to generation based on their profession and name may not be constrained to smiths and tailors alone. In medieval Germany, other professions such as miller, baker, shoemaker and webber were also organized in guilds and they also can be associated with certain physical characteristics like weight and stature. Therefore, the above mentioned ideas may also apply to those professions. Based on their data Bäumler, Forstner, and Stemmler (2003) constructed two groups of names, one group (Group I) was associated with a heavy weight and a strong physique, including names such as Smith, Miller, and Baker; the other group (Group II) was associated with a light weight and lean physique, including names such as Tailor, Shoemaker and Webber. Again, the successful athletes were cross-classified with respect to their respective disciplines in a 3 by 2 table (three rows: Group I, Group II and other names; two columns: heavy-and-medium and light).

Table 3:
Frequencies of successful athletes according to grouped last names 'Smith, Miller, Baker', 'Tailor, Shoemaker,Webber' and all other names by two categories of track-and-field (heavy-and-medium and light)

	Physical Statures				
	heavy-and-medium		light		
Successful athletes	f(o)	f(e)	f(o)	f(e)	Column marginals: $\sum$
Group I (Smiths, Miller, Baker)	a = 131	115.10	b = 78	93.90	A = 209
Group II (Tailors, Shoemaker, Webber)	c = 45	58.93	d = 62	48.07	B = 107
Other last names	6335	6336.98	5172	5170.02	11507
Row marginals: $\sum$	C = 6511		D = 5312		N = 11823

Note. Global $\chi^2 = 12.22$, df = 2, $p < .01$. f(o) = observed frequencies; f(e) = expected frequencies. The letters refer to Kimball's formula.

The corresponding global chi-square was significant $\chi^2 = 12.22$, df = 2, $p < .05$ suggesting a significant relationship between the three groups of athletes' names and the two branches of athletics:

$$\chi^2 = \sum_{i=1}^{k} \frac{(f(o) - f(e))^2}{f(e)} =$$

$$= \frac{(131-115.10)^2}{115.10} + \frac{(78-93.90)^2}{93.90} + \frac{(45-58.93)^2}{58.93} +$$

$$+ \frac{(62-48.08)^2}{48.08} + \frac{(6335-6336.98)^2}{6336.98} + \frac{(5172-5170.02)^2}{5170.02} =$$

$= 12.22$ with df = 2, $p < 0.05$.

Here, we also tested the twofold assumption that heavy-physique-names (i.e., Group I) are more successful in the heavy-and-medium disciplines and that light-physique names (i.e., Group II) are more successful in the light disciplines. Kimball's formula was applied and the two assumptions were supported by a significant chi-square:

$$\chi^2 = \frac{[A(Cd - Dc) - B(Cb - Da)]^2}{\frac{ABCD(A+B)(C+D)}{N}} =$$

$$= \frac{[209(6511 \times 62 - 5312 \times 45) - 107(6511 \times 78 - 5312 \times 131)]^2}{\frac{209 \times 107 \times 6511 \times 5312\,(209+107)\,(6511+5312)}{11823}} =$$

$= 4.9$ with df = 1, $p < 0.05$.

In sum, the above analyses show a strong relationship between the athletes names and their success in heavy-and-medium or light discipline.

5. Discussion

The analyses supported the hypothesis derived from a genetic-social theory that athletes with the names of 'Smith' and 'Tailor' were successful in different track-and-field disciplines and that this assortation of names and disciplines is based on a different physique of the name-bearers which is handed down from generation to generation through a combination of genetic and cultural laws. Whereas this association still holds for 'Tailors' and 'Smiths', it is much stronger for the 'Tailors', according to our first sample. This finding can be interpreted in two ways:

(1) There is a differential regression of the particular body-stature types. It seems that for the 'Tailors', who are presently more suited for and more successful in the light-stature athletic disciplines, a statistical regression to the mean happened more slowly than for the 'Smiths', who still lean toward the more heavy disciplines, but less strongly. This result might have psychological reasons: Men, who are typically built like a Smith, that is who are strongly built, have the highest prestige. 'Tailors' possessed less favorable physical characteristics and therefore less prestige. Consequently, 'Tailors' were less able than 'Smiths' to marry tall and strongly built women, because small men and tall women do not fit very well together, at least in the opinion of the society; a fact leading to the 'conservation' of a tailor-type physique. Contrary, 'Smiths' could afford to marry also small and less strongly built women, contributing to the somewhat 'dissolution' of a smithonian physique.

(2) There is a different biomechanical compatibility between the track-and-field disciplines and the specific occupational types. This interpretation can be seen as a supplement to the previous one. It is also feasible that the biomechanical requirements of the track-and-field disciplines involved demonstrate a higher compatibility for the typical physique of (black)smiths as compared to those of the tailors. That is, the biomechanical requirements of the 16 track-and-field disciplines would be less of a hindrance for individuals with the build of a (black)smith than for those with the stature of a tailor. The larger body-weight of the Smiths which can be seen as a hindrance for their participation in long-distance races, for example, can be compensated, at least in part, by the increased height - and therefore length of the legs - and by the cross-section of the muscles (strength) so that individuals with the physique of a (black)smith may also obtain reasonable achievements in long-distance races as well as in high-jumping. Individuals with a tailor-like physique, in contrast, are less able to make use of such a compensation for the ‘heavy’ track-and-field disciplines (e.g., hammer throw, shotput) since the required body mass is lacking in athletes with this type of physique. The various biomechanical adaptations to the occupational types required by (black)smiths and tailors and their suitability for track-and-field athletics in general can consequently be the basis for the different ‘conciseness’ of the both groups of names in the different athletic categories.

Further analyses provided evidence that the findings based on Smiths and Tailors are still holding true in todays society. Moreover, it seems to be that the underlying assumptions of body build and athletic success can be extended to larger groups of names which additionally include the names miller and baker on the onehand and the names webber and shoemaker on the other.

References

1. Bäumler, G. (1980). Differences in physique in men called ‘Smith’ or Tailor’ considered as results of a genetic effect dating back over several centuries. Personality and Individual Differences, 1, 308-310.
2. Bäumler, G. (1984). Unterschiede im Auftreten der Familiennamen ‘Schmied’ und ‘Schneider’ bei Spitzensportlern der Leichtathletik: Ein Beitrag zur epidemiologischen Humangenetik. Psychologische Beiträge, 26, 553-559.
3. Bäumler, G. and Lienert, G. A. (1987). KFA-Tests zur genetisch-sozialen Theorie der Sortierung von Familiennamen auf Körperbautypen. Psychologische Beiträge, 29, 587-595.
4. Bäumler, G., Forstner, M., and Stemmler, M. (2003). On the distribution of craftsmen family names in different categories of track-and-field athletics. A cross validation study. Sportonomics, 9, in press.
5. von Eye, A. (1990). Introduction to configural frequency analysis. The search for types and anti-types in cross-classification. Cambridge: Cambridge University Press.
6. von Eye, A. (2002). Configural Frequency Analysis - methods, models, and applications. Mahwah, NJ: Lawrence Erlbaum Publishers.
7. von Eye, A., Spiel, C., and Wood, P. K. (1996). Configural frequency analysis in applied psychological research. Applied Psychology, 4, 301-327.
8. Kimball, A. W. (1954). Short-cut formulae for the exact partition of chi-square in contingency tables. Biometrics, 10, 452-458.
9. Lautsch, E. and von Weber, S. (1995). Methoden und Anwendungen der Konfigurationsfrequenzanalyse (KFA). Weinheim: Beltz Psychologie Verlags Union.
10. Lienert, G. A. and Krauth, J. (1975). Configural frequency analysis as a statistical tool for detecting types. Educational and Psychological Measurement, 35, 231-238.

11. Lienert, G. A. and Netter, P. (1987). Nonparametric analysis of treatment-response tables by bipredictive configural frequency analysis. Methods of Information in Medicine, 26, 89-92.
12. Müller, M. J. and Stemmler, M. (2000). Bipredictive Configural Frequency Analysis of small sample MANOVA designs. Psychologische Beiträge, 42, 327-336.
13. Stemmler, M. (1994). A nonparametrical evaluation of ANOVA and MANOVA designs using interaction structure analysis. Biometrical Journal, 36, 911-925.