Scientometrics (2013) 94:523–539
DOI 10.1007/s11192-012-0760-3

How does scientific success relate to individual and organizational characteristics? A scientometric study of psychology researchers in the German-speaking countries

Hans P. W. Bauer · Gabriel Schui · Alexander von Eye · Günter Krampen

Received: 6 February 2012 / Published online: 3 May 2012

Abstract *Purpose*: To provide up-to-date bibliometric reference data describing the output and success of psychology researchers in the German-speaking countries, including lifetime publication and citation numbers, and to investigate associations of bibliometric measures with academic status and gender as well as the department characteristics of size and quota of senior researchers. *Method*: Queried literature databases using an extensive online register of academic psychologists in the German-speaking countries, obtaining valid data for 85 % ($N = 1742$) of the population of interest. *Findings*: Skewed distributions for publications and citations; maximum number of German-language (=native) publications much higher than maximum number of English-language publications; relatively large part of population publishing almost exclusively in German; publication count predictable by academic status, gender, department size, and quota of senior researchers; citation count predictable by publication count, status, department size, and quota of senior researchers; department characteristics interact with individual characteristics to produce specific conditions under which publication count and citation count are higher or lower than expected: combination of female gender, small department size and large quota of senior researchers is associated with particularly increased publication count; female gender and large department size are associated with decreased publication count; high publication count, large department size and low quota of senior researchers are associated with increased citation count; low publication count and large quota of senior researchers are associated with decreased citation count. *Conclusions*: Reference values for scientific output provided in this study provide an anchor for monitoring and international comparison; despite considerable noise in data, we show that interactions of individual and organizational characteristics are relevant for scientific success and should be investigated further, e.g. by adopting various measures of organizational diversity and tracing a population longitudinally.

H. P. W. Bauer (✉) · G. Schui · G. Krampen
Leibniz-Institute for Psychology Information (ZPID), University of Trier, 54286 Trier, Germany
e-mail: hans.bauer@zpid.de

A. von Eye
Department of Psychology, Michigan State University, 316 Psychology Building, East Lansing, MI 48824-1116, USA

Keywords German-speaking countries · Research output · Psychologists · Department characteristics · Academic status · Gender

Mathematics Subject Classification 62P25

JEL Classification I23

Introduction

Science policy and academic environments make heavy use of assessments of scientific performance, regarding various parameters (e.g., funding, tenure, paper publication) as well as dimensions (individual, institutional, national). Expert judgment is an important part of these assessments. However, it is resource-intensive and, therefore, increasingly complemented—sometimes even substituted—by scientometric indicators that seem easily obtained and suggest quantitative precision. When using such methods, objective factors must be taken into account that affect performance (as measured, for example, in units of output), and, consequently, the outcome of assessment. Objectivity is important not only because the validity of expert judgements might be compromised, but also because assessment results might not be appreciated by the stakeholders involved, or, more plainly, because the assessment might be regarded as unfair (Aksnes and Rip 2009; Pontille and Torny 2010).

Two factors that need to be accounted for in most cases are scientific discipline and country of residence.[1] There are differences in scientometric indices between disciplines (Garfield 1979), which, to a great degree, reflect academic "cultures" (Becher and Trowler 2001) that are usually not meant to be evaluated. In fact, in bibliometrics, discipline-normalized indices are routinely used when comparing different disciplines (Leydesdorff and Shin 2011; Schubert and Braun 1996).Similarly, there are significant differences between countries, as comparative studies show (May 1997; for psychology, see e.g. Navarrete-Cortes et al. 2010), which might be due to biases in the source database used (Bauserman 1997; Gibbs 1995; van Raan et al. 2011).

Therefore, it is important to have available up-to-date objective reference data on the relevant target populations, both to compare individual measures to population averages and to compare individual measures from different populations. This article examines the specific case of psychology in the *German-speaking countries (GSC: Germany, Austria, and parts of Switzerland)* by providing bibliometric measures (cumulative, i.e., life time publication and citation counts) of the population of psychologists working in research settings in the GSC in 2010. This information can also serve as a point of comparison with later studies of the same population and as a data unit for meta-analyses. Since lifetime measures are highly affected by time spent in the academic setting and academic status is highly correlated with the latter, academic status is taken into account, and results are reported separately for junior and senior researchers.

A second aim of this study is to identify some of the conditions that are associated with increased or decreased publication and citation success. To this end, researchers are profiled based on the individual characteristics of gender and academic status, as well as characteristics that describe the work environment: Department size and staff composition

[1] Relevant factors of course depend on the purpose of the assessment, which might just be to compare countries or disciplines.

in terms of numbers of junior and senior researchers. The effects of department size on research performance have been investigated several times, but results have been mixed (von Tunzelmann et al. 2003). Staff composition does not seem to have been investigated systematically in this context (for examples, see Moed et al. 1998; Wolszczak-Derlacz and Parteka 2011).Also, research on interactive effects of individual and environmental variables on scientific performance has been scarce (Horta and Lacy 2011). By identifying specific conditions associated with increased or decreased publication and citation success, we aim to provide initial evidence.

Associations of organizational as well as individual characteristics with research output are analyzed using a path model. Following up on significant interaction findings of the path model, statistical techniques focusing on single cells or profiles are then used to gain insight into the precise conditions associated with research output. This high resolution in statistical analyses is possible because the sample under study is large. In the following sections, we describe data acquisition and present variable definitions.

Method

Data for this study were gathered in three steps: 1. Names of researchers working at Psychology departments of Universities or publically funded research institutes in the GSC, along with information on their institutional affiliation and their academic titles, are drawn from a publically accessible online register. This constitutes the population of active scientists in academic psychology research contexts in the GSC. 2. Bibliometric data are obtained for this population by querying the online databases PSYNDEX und Social Science Citation Index/Science Citation Index (via Web of Science). 3. In a final step, data are cleaned to eliminate errors induced by shortcomings of bibliometric data collection procedures.

The population under study is constituted by the individuals that, in the fall of 2010, were listed in the"Hogrefe Psychologie-Kalender," which contains a complete directory of active psychology researchers in the GSC. Our data include psychologists holding at least a doctoral degree and working at universities offering a *Diplom*/Masters Program in Psychology. This information is used for an annually conducted monitoring ("ZPID-Monitor") on the internationalization of psychology from the GSC (see Krampen et al. 2005, 2011a) and was also employed in this study. The population consists of a total of 2,134 psychology researchers. 36 query names (1.7 % of total population) were deleted, because they are identical, and results would be confounded.

For each researcher, the following variables are coded, based on the "Psychologie-Kalender": Gender, academic status (holding a doctoral degree only or also having passed the postdoctoral *Habilitation*, which entitles one to be appointed as full professor; these two classes are labeled *junior* and *senior researchers*, respectively), and institutional affiliation (research department). In addition, the numbers of junior, senior, and total researchers are counted for each department. The total number of researchers is used as proxy for *department size* in this study. From the numbers of senior and junior researchers a measure of "seniority" in staff composition is computed by dividing the former through the latter (*department seniority*).Department size and department seniority were dichotomized (median split) for uses in nonparametric analyses. All department-level variables are then assigned to the individual records.

The next step in the creation of data involved querying two literature databases using researchers' names to obtain data for further filtering and for computation of scientometric

indices: The Web of Science databases SCI and SSCI (collectively referred to as "WoS" in the following) for publication and citation data,[2] and PSYNDEX for publication data. Because PSYNDEX's scope is limited to publications coauthored by psychologists based in the GSC, and because it extends to other types of publications besides journal articles and meeting abstracts (e.g., monographs, editions, dissertations), we consider it to be the most comprehensive as well as precise source concerning publication data for psychologists in the GSC. All data were collected between August and October of 2011.

The details of the database queries are as follows: In PSYNDEX, for each researcher, number of publications containing the researcher's name (last name and initials) in their author information were queried. No time limit was imposed upon the search; however, since the earliest publication year documented in PSYNDEX is 1977, the resulting index constitutes number of publications cumulated from 1977 up to 2011. The procedure was repeated with the constraint of English as publication language.

In the WoS, a similar name-based publication search was conducted. Results were limited to original articles (i.e., no editorials, letters, abstracts etc.) published since 1977. The WoS was then queried for the cumulative number of citations to each researcher's oeuvre. This was done by conducting a "cited reference search" for the author's last name and initials in the Web of Knowledge web interface. No limits on year of publications were imposed on the search. However, all self-citations were excluded from the results with an additional limiting query. Also, results were restricted to those citations coming from journals with WoS subject type "Psychology" and its subtypes. This was done because, other than PSYNDEX, WoS is not confined to a single discipline or to authors in a certain geographic region. Therefore, there is a much higher risk of false positive citations (due to author names being confounded). Limiting the results to those citations coming from psychology-related journals greatly reduces this risk, even though it may exclude some valid citations.

In addition to the *total number of citations*, we also obtained the number of citations by papers written in English. As a measure of *internationality*, the percentages of English-language publications in PSYNDEX and the percentages of citations by English-language articles in WoS were computed for each researcher.

To increase the validity of the data, two more cleaning steps were conducted:

- Individuals with no publications in PSYNDEX ($N = 89$; 4.2 %) were removed under the assumption that they are either at the very beginning of their career, mainly engaged in teaching, or not publishing for other reasons.
- To exclude individuals with a high probability of confounded author names (see Smalheiser and Torvik 2009, for a discussion on author-name disambiguation), we removed all persons whose publication count was higher in WoS than in PSYNDEX. In PSYNDEX, which we consider to be most complete and reliable source for publications in psychology from the GSC, the person with the highest publication count registers 920 publications, so a higher publication count in WoS seems highly implausible and suggests that the hits refer to publications of several persons of the same surname and initial(s), thus inflating citation results for the person in question. A total of 267 researchers (12.5 %) were removed in this step. Thus, 1,742 cases with reasonably high data quality remain. Since number of citations was zero for 133

[2] WoS publication data were only used for data cleaning (see below); publication data reported in the results section exclusively refer to PSYNDEX data.

researchers (6.2 %), the percentage of "international" citations could only be calculated for 1,609 psychology researchers (75.4 % of the population under study).

Statistical analyses of the data refer to a report of the descriptive characteristics of the population and—next—a log-linear path model of the demographic and bibliometric measures as well as analyses of interaction effects found in the model, the latter using Configural Frequency Analysis (CFA) (Lienert and Krauth 1975; von Eye and Gutiérrez Peña 2004; von Eye et al. 2010), which locates single cells in cross-tabulated frequency data that deviate from an expected pattern, thereby identifying specific conditions that affect output measures.

The main reasons for using this methodology are as follows. First, CFA is a method that allows one to identify "local associations" (Havránek and Lienert 1984) such as specific conditions associated with scientific success, which are of interest here. In cross-classifications of categorical variables, associations do not necessarily include all categories of the variables under study. Instead, selections of categories can be associated. These local associations manifest in the form of CFA types or antitypes. Methods of continuous variable analysis for the detection of local associations do not exist, in particular in the context of specific, hypotheses-defining statistical base models. Therefore, if local associations are of interest, CFA is the method of choice. Methods such as (hybrid) latent class analysis, Poisson regression, or generalized multilevel models serve to answer different questions than CFA. In addition, these methods are variable-oriented. In contrast, CFA is person-oriented (von Eye and Bergman 2003).

Second, CFA is a nonparametric statistical method. It does not require normal or multinormal variable distributions. Considering that the bibliometric variables in the present study are not normally distributed, robust or nonparametric methods are needed for analysis. Therefore, again, CFA can be considered a method most suitable for the present data situation.

Finally, the CFA base model that is used to estimate expected cell frequencies takes measurement dependency into account. Therefore, types and antitypes are sensitive to measurement dependencies only when the researchers intend to define types and antitypes on the basis of such dependency, and specify the base model accordingly.

Results

Data are analyzed on the individual as well as the department levels. A brief overview of the distributions of the department-level variables department size and department seniority across the departments examined is given in Table 1. Organizational units are comprised of an average of 38 researchers (range: 14–94).Seniority is quite widely spread, reaching from one senior researcher per four junior members to the exact opposite, i.e., four senior researchers per only one junior researcher.

Most researchers and departments in our population are located in Germany (about 80 %); Austria and Switzerland are represented with approx. 10 % each. Although the output measures that describe the researchers in these countries may differ, we consider the three countries sufficiently similar on a socio-cultural level to omit the country of residence as a variable in the following analyses.

Figure 1 displays the distribution of researchers by gender and academic status, as well as the frequencies that would be expected in case of stochastic independence. There are slightly more senior than junior researchers in the examined population, and there are more

Table 1 Descriptive summary statistics for department-level variables on department-level

	Department size	Department seniority
Valid *N*	56	56
Min	14	0.25
Max	94	4.00
Mean	38.11	1.26
SD	17.52	0.63
Skewness	0.86	1.69
Kurtosis	0.71	5.51
5th perc.	15	0.51
25th perc.	25	0.83
Median	36	1.10
75th perc.	50	1.57
95th perc.	71	2.23

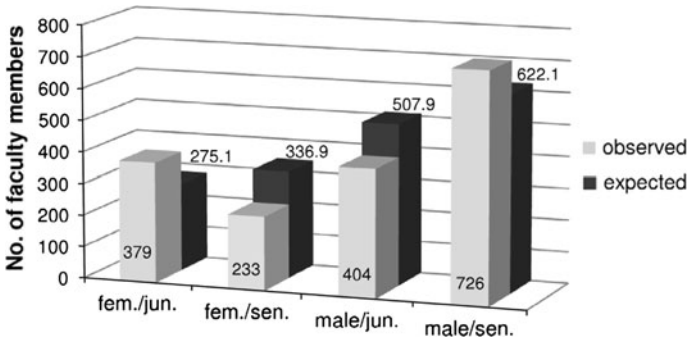

Fig. 1 Number of researchers in the observed population, according to gender and academic status

male than female researchers. In addition, male researchers hold disproportionately more often senior status. This discrepancy has been noted across nations and scientific disciplines and is generally regarded as a structural problem of academic environments (European Commission 1999; Ruest-Archambault 2008). Because of the resulting confound of gender and academic status (the latter being strongly correlated with output measures), and because the subject matter cannot be treated with adequate sophistication within space constraints, we chose not to differentiate between the gender groups in the description of the population with regards to output measures.

Details on output measure distributions are presented in Table 2. The scientometric variables are strongly skewed to the right and resemble Pareto probability functions. The distribution inequality is more marked for total citations than for total publications. Note that this becomes apparent only when looking at the percentile values provided; the skewness statistic is in fact higher for total publications, but this is because of very few extreme outliers in this variable. These outliers are due to the very high number of German-language publications in some cases; the upper bound appears to be much higher for German-language than for English-language publications. Because of the strong skewness, it is advisable to use percentile values in interpretation of these count variables, although mean and standard deviations are provided.

Table 2 Descriptive summary statistics for the population of interest and subpopulations according to academic status

		Valid N	Min	Max	Mean	SD	Skew	Kurt.	5th %	25th %	Median	75th %	95th %
Total publications	Senior	959	1	920	54.91	59.08	4.90	52.35	6	21	38	70	159
	Junior	783	1	125	12.81	13.06	3.26	18.84	1	4	9	17	36
	Total	1,742	1	920	35.99	49.36	5.50	67.22	2	8	21	44	127
English-language publications	Senior	959	0	205	19.77	26.27	2.63	9.37	0	3	11	26	74
	Junior	783	0	114	5.44	7.67	4.89	52.87	0	1	3	8	19
	Total	1,742	0	205	13.33	21.38	3.47	16.50	0	1	6	16	52
Non-English-language publications	Senior	959	0	865	35.14	47.50	7.19	102.38	2	10	22	43	113
	Junior	783	0	114	7.38	10.41	4.24	30.48	0	1	4	9	24
	Total	1,742	0	865	22.66	38.48	8.24	141.47	0	3	11	28	86
% English-language publications	Senior	959	0	100	35	30	0.54	−0.90	0	9	29	58	90
	Junior	783	0	100	42	37	0.30	−1.39	0	5	35	75	100
	Total	1,742	0	100	38	33	0.46	−1.11	0	8	31	67	100
Total citations	Senior	959	0	2,743	206.40	323.94	3.63	17.74	3	29	87	241	733
	Junior	783	0	1,362	44.45	99.34	6.78	64.61	0	3	15	45	166
	Total	1,742	0	2,743	133.61	262.05	4.56	28.62	0	9	41	139	580
Citations from English-language articles	Senior	959	0	2,606	167.45	295.94	3.91	20.52	0	12	57	194	676
	Junior	783	0	1,255	39.14	93.70	6.79	63.47	0	1	11	39	158
	Total	1,742	0	2,606	109.78	237.09	4.93	33.26	0	4	26	104	503
Citations from non-English-language articles	Senior	959	0	677	38.95	63.84	4.24	27.85	0	6	16	44	167
	Junior	783	0	199	5.31	12.46	7.41	87.41	0	0	1	6	23
	Total	1,742	0	677	23.83	50.91	5.38	44.88	0	1	6	23	110
% citations from English-language articles	Senior	935	0	100	67	30	−0.74	−0.72	8	45	78	93	99
	Junior	675	0	100	77	31	−1.39	0.64	0	66	93	100	100
	Total	1,610	0	100	71	31	−0.97	−0.33	0	50	85	96	100

It is interesting to note that, in the light of recent debates in the psychology community of the GSC about anglicization of the discipline (Gigerenzer et al. 1999; Tack 1994), the majority of the psychology publications from the GSC is still published in German. About 25 % of the population almost exclusively publishes in German. In contrast, about 5 % solely publish in English.

It must be noted that, although publication data are likely to be representative of the full spectrum of publications in the population, citation data only reflect the impact in journal articles, meeting abstracts, and proceedings papers, because these are the only document types indexed in the WoS. Even as books have recently found their way into the WoS, their numbers are still low and the selection is limited. In addition, the WoS counts citations of books only for the first two authors/editors. Interpretation is even more restricted in case of the output variables "Non-/English-language citations" and "Percentage of citations from English articles", since English-language articles are vastly overrepresented in comparison to other languages in the WoS (in Psychology, the quota of English-language records has been just shy of 95 % over the last decade, which is unlikely to reflect the true proportion; see Krampen et al. 2011a). Thus, results should be interpreted only in comparison to results also derived from the WoS.

In the following sections, total citation and total publication data are analyzed using log-linear path modeling and CFA to examine the amount of variability explained in research output and citation counts, and to identify conditions under which significantly higher or lower output values occur than would be expected. These methods are appropriate, because the distributions of the publication and citation measures are extremely skewed and leptokurtic. Log-linear path models and CFA analyze frequency data from cross-tabulated categorical variables. Due to the large population size, this can be done with no problems. To obtain a full-count cross-tabulation, median splits of the non-binary variables for the following analyses are used.

Three models were estimated for the cross-classification of the following six variables: Gender (G; 1 = male, 2 = female), academic status (T; 1 = senior, 2 = junior), department size (Z; 1 = below average, 2 = above average; dichotomized), department seniority (S; 1 = below average, 2 = above average; dichotomized ratio), total number of publications (P; 1 = below average, 2 = above average; dichotomized frequency), and total number of citations (C; 1 = below average, 2 = above average; dichotomized frequency).

Three models are estimated. Each of these models is a path model, that is, a model with a conditional probability structure (Goodman 1973; Vermunt 1997; von Eye and Mun 2012). Each of the models predicts, firstly, the number of publications (P) from the four variables that describe the respondents (G, T) and, secondly, the departments they work in (Z, S), and the number of citations (C) from the number of publications. The three models are, in conditional probability notation,

$$\pi_{GTZSPC} = \pi_{GTZS}\pi_{P|GTZS}\pi_{C|GTZSP}, \tag{1}$$

$$\pi_{GTZSPC} = \pi_{GTZS}\pi_{P|GTZS}\pi_{C|TZSP}, \tag{2}$$

and

$$\pi_{GTZSPC} = \pi_{GTZS}\pi_{P|TZS}\pi_{C|TZSP}. \tag{3}$$

The first of these models (Model 1 in Table 3) is saturated. We use it as a reference model for the following two. The second model (Model 2) removes the Gender effect on the citation record. Thus, this model does consider G, T, Z, and S as predictors of the number of published works, but it only considers T, Z, S, and P as predictors of the number

Table 3 Model fit statistics for conditional probability models

Model[a]	LR-X^2	d*f*	*p*
1	0.00	0	–
2	15.90	16	0.46
3	46.46	24	<0.01

[a] Model 1: saturated model, Model 2: gender effect removed from prediction of citations, Model 3: gender effect removed from prediction of citations and publications

of citations. The third model (Model 3) removes Gender entirely from consideration as a predictor of publication and citation numbers. Table 3 displays overall goodness-of-fit information for these three models. All models were estimated using Lem (Vermunt 1993).

Table 3 shows that Model 2 is the only one that can be retained. It comes with no significant model-data discrepancies, keeps the Gender effect on publication frequencies, but does not propose that Gender plays a role in citation frequencies. The more parsimonious Model 3 which removes Gender from all predictive relationships of the model does come with significant model-data discrepancies. Therefore, we retain and interpret Model 2.

In all, the prediction of publication numbers from variables that describe the respondents and the departments they work in is significant and solid. The classification error-based R^2 equivalent for this part of the path model suggests that 55.6 % of the variability of the publication numbers can be explained from the four predictors.

Of the parameters that represent the links of the variables that describe the respondents and the departments they are working in with publication numbers, the one that represents the highest-order interaction, [G, T, Z, S, P], is significant. We obtain the estimate $b = 0.08$, with $se = 0.04$ and $p = 0.014$. This significant term suggests that the joint frequency distribution of G, T, Z, and S differs over the categories of P. To explain this rather complex relationship (a 5-way interaction) we exploit the formal similarities of logistic regression and Prediction CFA (P-CFA; von Eye et al. 2005). The five-way interaction [G, T, Z, S, P] in the present path model is equivalent to the regression of P onto the four-way joint distribution of G, T, Z, and S, in a standard logistic regression model that includes the four-way interaction among the four predictors. The significance tests are exactly the same. As von Eye et al. (2005) showed, the underlying model is also identical to the base model of P-CFA. Therefore, we employ P-CFA to identify those patterns of G, T, Z, and S that differ in particular over the two levels of P. Put another way, reported effects refer to comparisons of frequencies of cases with low versus high publication numbers within a given pattern of values of variables gender, status, department size, and department seniority (effects are thus directional). Such a pattern is referred to as a "profile" in the following. The results of P-CFA are summarized in Table 4.

P-CFA suggests that, with the exception of the first profile, 1111, frequencies in all profiles differ significantly over the two levels of P. Notably, when comparing any profile where gender is female with the corresponding profile where gender is male, differences are always in the same direction. That is, any combination of T, Z and S values that contributes to low or high publication numbers of female researchers also does so for male researchers. However, effects are in general more pronounced in females, bringing about the full 5-way-interaction.

Due to space constraints, we only interpret the discrepancies that are most extreme in units of the z test used to test the null hypothesis of no differences. The most extreme

Table 4 Results of P-CFA for total publications

Profile[a]	Frequency	z	$p(z)$	Type
1111_1	10	−0.17	0.4318	–
1111_2	11			
1112_1	9	−6.33	0.0000	Discrimination type
1112_2	62			
1121_1	42	4.77	0.0000	Discrimination type
1121_2	9			
1122_1	42	3.55	0.0002	Discrimination type
1122_2	16			
1211_1	18	−3.07	0.0011	Discrimination type
1211_2	42			
1212_1	38	−7.69	0.0000	Discrimination type
1212_2	136			
1221_1	52	6.04	0.0000	Discrimination type
1221_2	7			
1222_1	51	3.56	0.0002	Discrimination type
1222_2	22			
2111_1	20	−4.13	0.0000	Discrimination type
2111_2	56			
2112_1	44	−10.19	0.0000	Discrimination type
2112_2	192			
2121_1	142	10.09	0.0000	Discrimination type
2121_2	21			
2122_1	151	8.29	0.0000	Discrimination type
2122_2	44			
2211_1	27	−2.49	0.0064	Discrimination type
2211_2	49			
2212_1	65	−7.75	0.0000	Discrimination type
2212_2	180			
2221_1	88	7.13	0.0000	Discrimination type
2221_2	18			
2222_1	63	5.66	0.0000	Discrimination type
2222_2	15			

[a] Profiles denote the values the variables take by order and numbering; Order of and value labels of variables are: Gender (G; 1 = male, 2 = female), Status (T; 1 = senior, 2 = junior), Department size (Z; 1 = small, 2 = large), Department seniority (S; 1 = small quota of seniors, 2 = large quota) and (separated by underscore) number of publications (P; 1 = few, 2 = many; predicted variable); significance tests refer to differences in frequencies of a given GTZS profile when P = 1 versus P = 2

z score was calculated for the difference between 2112_1 and 2112_2, where the order of variables is G, T, Z, S, and P. 44 researchers with below average publication numbers are female, have senior status, work in small departments with above average numbers of senior researchers. 192 researchers show the same profile but exhibit higher publication counts than average. In other words, profile 2112 allows one to predict above average publication numbers. In contrast, profile 2121 (female, senior, large department size; relatively few senior researchers) allows one to predict below average publication numbers. This is the second-most extreme discrepancy. The third-most extreme discrepancy was found for profile 2122 (female, senior, small department size, relatively many senior researchers). This profile also allows one to predict below average publication numbers.

The two profiles that are predictive of below average publication numbers differ only in the last variable. Therefore, a theorem by Quine and McCluskey can be applied (cf. Hoernes and Heilweil 1964; von Eye and Brandtstädter 1982), and the two profiles can be fused to form the aggregate profile 212#, where the number sign (hash) indicates that regardless of the relative number of senior researchers in a department, the profile "female, senior, and large department size" is predictive of below average publication numbers.

The three previously described effects all referred to female researchers. As mentioned, effects go in the same direction for male researchers, but are much less pronounced. This is not the case for the fourth most extreme effect, that is, the one for profile 2212. It is only negligibly larger than that for its "male" counterpart, 1212. In both cases, the combination of junior status, small department, and large quota of senior researchers (profile #212) makes a higher publication count more likely.

A number of the lower-order interactions of this part of the prediction structure were significant also. However, each of these is a lower-order relative of the interaction just interpreted. Therefore, there is no need for additional interpretation.

Moving on to citation counts, 55.3 % of their variability can be explained with the retained path model from the descriptive variables (excluding Gender) plus the publication numbers. The highest-order significant interaction of this part of the path model is the one among department size, department seniority, publication numbers, and citation numbers [Z, S, P, C]. For the parameter for this interaction, we obtain the estimate $b = -0.07$, with $se = 0.04$ and $p = 0.022$. This interaction suggests that the association between Z, S, and P differs across the two levels of citation success. For a more detailed interpretation we again perform a P-CFA in which we predict C from the patterns of Z, S, and P (Table not shown here).

P-CFA suggests that profile 111 is again the only one that fails to allow one to predict whether a researcher has above versus below average citation numbers. All other profiles do allow such prediction. Profile 121 is the most extreme. Based on this profile, it can be predicted that researchers who work in small departments with relatively many senior researchers and have below average numbers of publications will overwhelmingly be among those with small citation numbers. The same applies for researchers with Profile 221. Again, these two profiles, both predictive of low citation numbers, differ only in their indicator of size of department. Aggregation results in profile #21. In other words, working in an environment with many senior researchers and having few publications is predictive of small citation numbers, regardless of size of department (and gender). Finally, profile 212 is predictive of high citation numbers. Researchers who work in large departments with relatively few senior researchers and have above average numbers of publications will be cited frequently.

One three-way and two of the two-way interactions of this part of the path model are also significant. However, all of these are lower-order interactions of the four-way interaction interpreted in the last paragraph. Therefore, there is no need to interpret the lower-order interactions. None of the other four-way interactions was significant.

Discussion

Objectives of this study were twofold: first, we present in detail up to date-information about lifetime publication productivity and citation impact of the population of psychology researchers in the GSC (Austria, Germany, and parts of Switzerland). This allows one to evaluate research output relative to a 'reference standard' specific to psychology in the

GSC as well as comparisons of this standard to those in other nations and scientific disciplines and/or at other points in time. Second, we explore conditions associated with higher versus lower scientific productivity and impact, considering characteristics of both the individual researcher and the departments they are working in.

Our analyses refer to almost the entire population, excluding only cases for which acceptable scientometric data quality could not be ascertained. Since exclusion was mainly based on name frequencies, a variable that should be independent of any research performance measure, the results of our analyses have a good validity. Thus, it can be stated that on average lifetime publication count of psychology researchers in the GSC amounts to 36 (MD: 21) publications. A recent study by Duffy et al. (2011) investigated industrial and counseling psychology researchers in North America, with a mean lifetime publication score of 18.7 publications. Even when taking into account the differences between the samples (e.g., in nationality, first language, and scientific subdiscipline), the discrepancy to our results is quite striking. It is most probably, in part, due to the fact that the output measure used in our study includes non-article publications (e.g., books, book chapters, psychological tests, reports) too.

The total number of psychology publications has been rising steadily over the last decades in the GSC as well as elsewhere, and is likely to continue to increase (Krampen and Schui 2009; Krampen et al. 2011b; Krampen and Wiesenhütter 1993).This, however, could be due to an increasing number of researchers as well as to increasing activity levels. Looking at earlier implementations of the ZPID-Monitor, which uses the same population, it can be observed that the size of the population has been rising from about 1,500 researchers in the year 2001 (Krampen et al. 2005) to more than 2,000 in 2011 (Krampen et al. 2011a).

One must also be careful in comparing citation counts gleaned from Web of Science. WoS offers both the possibility to query the number of citations of an *author's works recorded in WoS* as well as the number of *all articles recorded in WoS that include the author's name in their reference list*. When name ambiguities can be appropriately dealt with, the latter option is to be preferred since it also includes citations of works not indexed in WoS (non-source items), such as books. We found that psychology researchers in the GSC have an average of 134 (MD: 41) citations per person (in WoS-recorded publications). This is a much lower number than the 329 citations per researcher reported by Duffy et al. (2011), pointing at the necessity of nation-specific standards (see also Navarrete-Cortes et al. 2010).

Looking at the distributions of the publication output measures, we found that the distribution of German-language publications is much more uneven than that of English-language publications. This may be due to the fact that the latter are often more difficult to produce for a non-native speaker (Cheung 2010; Montada and Krampen 2001) and therefore their number is likely to be limited by a ceiling effect.

As far as publication language is concerned, the "average" psychology researcher in the GSC publishes 60–65 % of his or her work in German and 35–40 % in English. This ratio is distributed more evenly than total publication numbers, i.e., there are some researchers that prefer to publish exclusively in German and others who do so in English. The former group is larger, which should not be surprising given the fact that psychology has its roots both in natural sciences as well as in humanities and social sciences, with the latter depending more strongly on culture- and language-specific factors (Schui and Krampen 2007). However, the quota of English-language publications from the GSC has been growing steadily and is expected to rise even further (Krampen et al. 2011a). This applies

particularly to subdisciplines oriented towards natural science and experimental paradigms, such as experimental and biological psychology.

Since our scientometric output variables are cumulated measures, it is important to provide separate standards for researchers of differing academic status, which is closely related to time spent in the academic setting. Unsurprisingly, median publications and citations are higher for senior than for junior researchers (38 vs. 9 and 87 vs. 15, respectively).However, the opposite is true for the mean (as well as median) quota of English-language publications (35 vs. 42 %). The most intuitive explanation is that younger researchers have been more strongly socialized towards publishing in English.

Research performance not only depends on individual characteristics like the ones just discussed, but is also associated with organizational context variables as well as the interaction of both (Horta and Lacy 2011). In the present investigation, academic status and gender on the one hand, and department size and department seniority (ratio of senior to junior researchers) on the other were analyzed with regard to lifetime publication and citation counts in two steps: first, fitting a log-linear path model to the data, and second, analyzing interaction effects identified with the path model in detail, using CFA.

More than 50 % of the variability in publication numbers was explained by individual and department characteristics in a log-linear path model. In turn, a similarly large part of the variability in citation numbers could be explained by academic status, department size, and department seniority in conjunction with publication numbers. Higher-order interaction effects contributed to variability explained in both cases, indicating that research performance is sensitive to quite specific conditions. These were followed up on with CFA.

Publication numbers were sensitive to the specific pattern of gender, academic status, department size, and department seniority. Patterns of academic status, department size, and department seniority were similar in direction, but not in magnitude of effect with regard to gender: Effects were in almost all cases more pronounced in female researchers, indicating that organizational structure might be more closely associated with their research output and career success. Increased publication output is strongly associated with a combination of female gender, senior academic status, small department size, and a large quota of senior researchers in the department. The same organizational context is also positively associated with publication output in junior researchers, this time to the same degree for male and female researchers. Possibly, a productive atmosphere is created by a smaller, less anonymous work environment with well acquainted senior researchers; from this atmosphere, junior researchers profit as well. On the other hand, senior status female researchers in large departments have a lower publication output than expected. A rather tentative explanation could be that increased competition for resources between or within departments might be a hindrance to career advancement of female researchers (see Niederle and Vesterlund 2007).

Interestingly, results for citation measures partly appear to go in the opposite direction. A CFA was conducted for profiles of department size, department seniority, and publication count, since this was the highest-order interaction identified in the path model. Publications of researchers with a high publication output who work in large departments with relatively few senior researchers attract more citations than expected. That is, given that a researcher publishes a lot, these organizational attributes contribute to a high citation count which are the exact opposite of those organizational variables that contribute to a large publication output. In contrast, a particularly low impact is found for publications of individuals having published less, but also working in a department with a large quota of senior researchers.

While the explanations concerning publication output assume some kind of effect of organizational variables upon the performance measure, the results for citations might in fact be best explained the other way around: "High impact researchers", already well known, might attract funding and therefore young researchers, leading to larger departments and a smaller quota of senior researchers. This would be in line with results of Moed et al. (1998), who found that, over the course of 10 years, scientific institutes with a high impact in the beginning of this time period displayed a particularly large rise in third party funding as well as in quota of junior researchers. In a similar vein, it might be that the low citation count of individuals with low publication numbers working in departments with many senior researchers is due to the fact that these individuals "accumulate" in possibly less-well funded departments, passing the *Habilitation* (which is the criterion for senior status in our investigation) at some time but do not advance much further beyond that point.

The interpretations provided above concerning the effects of organizational characteristics also highlight an important limitation of the present study: Researchers were sampled only at one point in time, in 2010, in a cross-sectional manner. The large sample thus obtained made the analysis of effects of very specific conditions (profiles of individual and organizational characteristics) possible. On the other hand, it cannot be ascertained whether success is precedent or antecedent to the 2010 institutional affiliation, and results should only be interpreted as associations. The main strength of the present analysis lies in the specificity of the conditions investigated. Associations of these conditions with scientific success should be followed up upon in longitudinal studies.

In any case, at least in psychology departments in the GSC, the organizational characteristics of department size and department seniority are associated with scientific performance as is their interaction with the individual characteristics of gender and academic status. It has been concluded that in science, organizational size exerts its effects primarily through the size of the immediate research team (von Tunzelmann et al. 2003).In our study, department size refers to organizational units of 14–94 researchers, organized into several chairs, specialized in certain subdisciplines, such as cognitive, social, or clinical psychology. Nevertheless, there are significant effects, likely owing to our large sample size. We assume that the variable "department size" likely acts as a proxy for team size.

Similarly, "department seniority" is only one of many facets of staff composition which might affect research output. Gender composition is another one (see, for example, Wolszczak-Derlacz and Parteka 2011) and so is ratio of permanently to temporarily employed staff or quota of foreign and external coworkers. More generally, research on scientific output can thus be placed in the context of research on the effects of organizational "diversity" (Jackson and Ruderman 1995).Still, even ignoring those other factors, one would expect the workplace dynamics to be different when there is one senior researcher per four junior researchers, as compared to when this ratio is reversed, as was the case in some of the departments investigated here.

The present investigation analyzed in detail characteristics of the distributions of two scientometric measures often used in evaluations of research performance (and consequently for research personnel and policy decisions), i.e., cumulative publication and citation counts, for psychology researchers in the German-speaking countries. This establishes specific parameter values which can be used as a reference standard for conducting further assessments and comparisons as well as for monitoring the development of the standards themselves. Also, a coarse differentiation was made according to academic status. In a more detailed inferential analysis, variability in publication and citation counts was explained through a log-linear path model. In an analysis of interaction effects

contributing to variance explained, specific conditions are identified which are associated with publication or citation success, respectively. Size of department and the ratio of junior to senior researchers are found to have an effect in conjunction with the individual characteristics of gender and academic status. However, especially with regard to the association of organizational variables and performance measures, it cannot be established whether the former influence the latter or whether the reverse is true. Also, the inclusion of additional explanatory variables, individual (e.g., subdiscipline of researcher) as well as organizational (e.g., team size and staff diversity) may help to clear up the structure of effects. Therefore, future research should be longitudinal, should consider–in addition–data on scientific subdisciplines, team size and diversity and should use publication output concurrent to the current working environment as well as time-lagged citations.

References

Aksnes, D. W., & Rip, A. (2009). Researchers' perceptions of citations. *Research Policy, 38*, 895–905. doi:10.1016/j.respol.2009.02.001.

Bauserman, R. (1997). International representation in the psychological literature. *International Journal of Psychology, 32*, 107–112. doi:10.1080/002075997400908.

Becher, T., & Trowler, P. R. (2001). *Academic tribes and territories: Intellectual enquiry and the culture of disciplines* (2nd ed.). Buckingham: The Society for Research into Higher Education & Open University Press.

Cheung, Y. L. (2010). First publications in refereed English journals: Difficulties, coping strategies, and recommendations for student training. *System, 38*, 134–141. doi:10.1016/j.system.2009.12.012.

Duffy, R. D., Jadidian, A., Webster, G. D., & Sandell, K. J. (2011). The research productivity of academic psychologists: Assessment, trends, and best practice recommendations. *Scientometrics, 89*, 207–227. doi:10.1007/s11192-011-0452-4.

European Commission. (1999). *Women and science: Mobilising women to enrich European research.*COM (2009) 447 final. Retrieved December 13, 2011, from http://cordis.europa.eu/improving/women/communication.htm.

Garfield, E. (1979). Is citation analysis a legitimate evaluation tool? *Scientometrics, 1*, 359–375. doi:10.1007/BF02019306.

Gibbs, W. W. (1995). Lost science in the third world. *Scientific American, 273*(2), 76–83.

Gigerenzer, G., Rösler, F., Spada, H., Amelang, M., Bierhoff, H. W., Ferstl, R., et al. (1999). Internationalisierung der psychologischen Forschung in Deutschland, Österreich und der Schweiz: Sieben Empfehlungen [Internationalisation of psychological research in Germany, Austria and Switzerland: Seven recommendations]. *Psychologische Rundschau, 50*, 101–105. doi:10.1026//0033-3042.50.2.101.

Goodman, L. A. (1973). Causal analysis of data from panel studies and other kinds of surveys. *American Journal of Sociology, 78*, 1135–1191.

Havránek, T., & Lienert, G. A. (1984). Local and regional versus global contingency testing. *Biometrical Journal, 26*, 483–494.

Hoernes, G. E., & Heilweil, M. F. (1964). *Introduction to Boolean algebra and logic design*. New York: McGraw-Hill.

Horta, H., & Lacy, T. A. (2011). How does size matter for science? Exploring the effects of research unit size on academics' scientific productivity and information exchange behaviors. *Science and Public Policy, 38*, 449–462. doi:10.3152/030234211X12960315267813.

Jackson, S. E., & Ruderman, M. N. (1995). *Diversity in work teams: Research paradigms for a changing workplace*. Washington, DC: American Psychological Association.

Krampen, G., & Schui, G. (2009). ZPID-Monitor 2007 zur Internationalität der Psychologie aus dem deutschsprachigen Bereich: Der Kurzbericht [ZPID-monitor 2007 on the internationality of psychology in the German-speaking countries: Short report]. *Psychologische Rundschau, 60*, 184–185. doi:10.1026/0033-3042.60.3.184.

Krampen, G., & Wiesenhütter, J. (1993). Bibliometrische Befunde zur Entwicklung der Teildisziplinen der Psychologie [Bibliometric findings on the development of subdisciplines of psychology]. *Psychologische Rundschau, 44*, 25–34.

Krampen, G., Montada, L., Müller, M. M., & Schui, G. (2005). *Internationalität und Internationalisierung der deutschsprachigen Psychologie [Internationality and internationalization of German-language psychology]*. Göttingen: Hogrefe.

Krampen, G., Schui, G., & Bauer, H. (2011a). ZPID-Monitor 2009 zur internationalität der Psychologie aus dem deutschsprachigen Bereich: Der Kurzbericht [ZPID-Monitor 2009 on the internationality of psychology in the German-speaking countries: Short report]. *Psychologische Rundschau, 62*, 244–248. doi:10.1026/0033-3042/a000094.

Krampen, G., von Eye, A., & Schui, G. (2011b). Forecasting trends of development of psychology from a bibliometric perspective. *Scientometrics, 87*, 687–694. doi:10.1007/s11192-011-0357-2.

Leydesdorff, L., & Shin, J. C. (2011). How to evaluate universities in terms of their relative citation impacts: Fractional counting of citations and the normalization of differences among disciplines. *Journal of the American Society for Information Science and Technology, 62*, 1146–1155. doi:10.1002/asi.21511.

Lienert, G. A., & Krauth, J. (1975). Configural frequency analysis as a statistical tool for defining types. *Educational and Psychological Measurement, 35*, 231–238.

May, R. M. (1997). The scientific wealth of nations. *Science, 275*, 793–796. doi:10.1126/science.275.5301.793.

Moed, H. F., Luwel, M., Houben, J. A., Spruyt, E., & Berghe, H. (1998). The effects of changes in the funding structure of the Flemish universities on their research capacity, productivity and impact during the 1980's and early 1990's. *Scientometrics, 43*, 231–255. doi:10.1007/BF02458409.

Montada, L., & Krampen, G. (2001). Internationalität und Internationalisierung der deutschsprachigen Psychologie [Internationality and internationalisation of psychology in the German-speaking countries]. In R. K. Silbereisen & D. Frey (Eds.), *Perspektiven der Psychologie. Einführung und Standortbestimmung* (pp. 282–300). Weinheim: Beltz.

Navarrete-Cortes, J., Fernandez-Lopez, J. A., Lopez-Baena, A., Quevedo-Blasco, R., & Buela-Casal, G. (2010). Global psychology: A bibliometric analysis of Web of Science publications. *Universitas Psychologica, 9*, 553–567.

Niederle, M., & Vesterlund, L. (2007). Do women shy away from competition? Do men compete too much? *The Quarterly Journal of Economics, 122*, 1067–1101. doi:10.1162/qjec.122.3.1067.

Pontille, D., & Torny, D. (2010). The controversial policies of journal ratings: evaluating social sciences and humanities. *Research Evaluation, 19*, 347–360. doi:10.3152/095820210X12809191250889.

Ruest-Archambault, E. (2008). *Benchmarking policy measures for gender equality in science*. Luxembourg: Office for Official Publications of the European Communities.

Schubert, A., & Braun, T. (1996). Cross-field normalization of scientometric indicators. *Scientometrics, 36*, 311–324. doi:10.1007/BF02129597.

Schui, G., & Krampen, G. (2007). Zur Internationalität der Pädagogischen Psychologie aus dem deutschsprachigen Bereich. Personen- und fachbezogene bibliometrische Publikations- und Zitationsanalysen [On the internationality of educational psychology in the German-speaking countries. Person- and discipline-centered bibliometric publication and citation analyses]. *Zeitschrift für Pädagogische Psychologie, 21*, 97–106.

Smalheiser, N. R., & Torvik, V. I. (2009). Author name disambiguation. *Annual Review of Information Science and Technology, 43*, 287–313.

Tack, W. H. (1994). Bericht über Reaktionen auf einen Bericht: Zu Keul, Gigerenzer und Stroebes SSCI-Analyse [reactions to Keul, Gigerenzer and Stroebe's SSCI analysis]. *Psychologische Rundschau, 45*, 108–111.

van Raan, A. F. J., Leeuwen, T. N., & Visser, M. S. (2011). Severe language effect in university rankings: particularly Germany and France are wronged in citation-based rankings. *Scientometrics, 88*, 495–498. doi:10.1007/s11192-011-0382-1.

Vermunt, J. K. (1997). *Log-linear models for event histories. Advanced quantitative techniques in the social sciences*. Thousand Oaks: Sage Publications.

von Eye, A., & Bergman, L. R. (2003). Research strategies in developmental psychopathology: Dimensional identity and the person-oriented approach. *Development and Psychopathology, 15*, 553–580.

von Eye, A., & Brandtstädter, J. (1982). Systematization of results of configuration frequency analysis by minimizing Boolean functions. In H. Caussinus, P. Ettinger, & J. R. Mathieu (Eds.), *COMPSTAT 1982, part II: Short communications, summaries of posters* (pp. 91–92). Wien: Physica.

von Eye, A., & Gutiérrez Peña, E. (2004). Configural frequency analysis: The search for extreme cells. *Journal of Applied Statistics, 31*, 981–997. doi:10.1080/0266476042000270545.

von Eye, A., & Mun, E. Y. (2012). *Log-linear modeling—Concepts, interpretation, and applications with SYSTAT, Lem, and R*. New York: Wiley. (in press).

von Eye, A., Mair, P., & Bogat, G. A. (2005). Prediction models for configural frequency analysis. *Psychology Science, 47*, 342–355.

von Eye, A., Mair, P., & Mun, E.-Y. (2010). *Advances in configural frequency analysis*. New York: Guilford Press.

von Tunzelmann, N., Ranga, M., Martin, B., & Geuna, A. (2003). *The effects of size on research performance: A SPRU Review*. Retrieved December 13, 2011, from http://www.sussex.ac.uk/Users/prff0/Publications/Size%20HE%20research.pdf.

Wolszczak-Derlacz, J., & Parteka, A. (2011). Efficiency of European public higher education institutions: a two-stage multicountry approach. *Scientometrics, 89*, 887–917. doi:10.1007/s11192-011-0484-9.