creative scientists,
star scientist,
interdisciplinarity....

Jointly published by Akadémiai Kiadó, Budapest and Springer, Dordrecht

Scientometrics, Vol. 70, No. 3 (2007) 811–830
DOI: 10.1007/s11192-007-0313-3

Characterizing creative scientists in nano-S&T: Productivity, multidisciplinarity, and network brokerage in a longitudinal perspective

THOMAS HEINZE,[a] GERRIT BAUER[b]

[a] *Fraunhofer Institute for Systems and Innovations Research, Karlsruhe (Germany)*
[b] *Mannheim Centre for European Social Research (MZES), Mannheim (Germany)*

creative scientist can be predicted on the basis of the total # of citations

While some believe that publication and citation scores are key predictors of breakthroughs in science, others claim that people who work at the intersection of scientific communities are more likely to be familiar with selecting and synthesizing alternatives into novel ideas. This paper contributes to this controversy by presenting a longitudinal comparison of highly creative scientists with equally productive researchers. The sample of creative scientists is identified by combining information on science awards and nominations by international peers covering research accomplishments in the mid-1990s. Results suggest that it is not only the sheer quantity of publications that causes scientists to produce creative pieces of work. Rather, their ability to effectively communicate with otherwise disconnected peers and to address a broader work spectrum also enhances their chances to be widely cited and to develop novel ideas.

Introduction

Creative capabilities are an important cornerstone of progress in science and technology, and a precondition for advances in other societal domains. However, creativity in scientific research has been given only limited attention in science studies. Our current knowledge about how unconventional, path-opening solutions in science emerge – and about how they can be fostered institutionally – is still rather incomplete.

Received November 15, 2006

Address for correspondence:
THOMAS HEINZE
Fraunhofer Institute for Systems and Innovations Research
Breslauer Strasse 48, 76139 Karlsruhe, Germany
E-mail: thomas.heinze@isi.fhg.de

0138–9130/US $ 20.00

Consequently, if we want to advance our understanding of the dynamics of science at research frontiers, we need to know more about what creative research accomplishments are, how they can be identified, in which organizations they occur most often, and what distinguishes highly creative scientists and groups from their peers.

This paper approaches scientific creativity at the level of individual scientists. Scientists with a record of highly creative research accomplishments in the field of nanoscience and nanotechnology (referred to as "nano-S&T") are compared with a matched comparison group of peer scientists. Examining their publication record, their citation patterns, the disciplinary scope of their published work, and the structure of their professional networks identifies those dimensions and characteristics which distinguish highly creative scientists from other researchers. Nano-S&T is a relatively young domain of scientific endeavor and embraces research areas such as applied physics, materials science, physical chemistry, physics of condensed matter, biochemistry and molecular biology, and polymer science and engineering (HULLMANN & MEYER, 2003; HEINZE, 2006).

Our analysis is informed by two influential yet unconnected sets of arguments in the literature on the emergence of scientific creativity and new ideas. The first argument is taken from Simonton who claims that prolific scientists have a higher probability of their work being selected as "creative" by their peer scientists. Publication and citation scores are believed to be important predictors of breakthroughs in science (SIMONTON, 1999; 2004). In contrast, in his analysis of the performance of company managers, Burt argues that people who live at the intersection of social groups are more likely to be familiar with selecting and synthesizing alternatives into novel ideas (BURT, 1992; 2004). Transferring these insights to the world of science implies that scientists who connect homogeneous groups, such as disciplinary communities or research fields, have a higher probability of exposure to alternative ways of thinking and behaving.

Starting with these arguments, we test the hypothesis that creative scientists in the field of nano-S&T can be predicted from their citation and publication record. We also test the claim that creative scientists, relative to their peers, belong to professional networks with access to richer and more diverse expertise, and that they address a broader disciplinary spectrum in their work. The sample of creative scientists is identified by combining information on science awards with nominations by international peers covering research accomplishments from the mid-1990s. Independent variables include the number of publications and citations, the size of co-authorship networks, information brokerage, and multidisciplinarity indices. We apply a longitudinal multi-method research design based on five consecutive periods of three years each, spanning the years 1990 to 2004.

The results suggest that while highly creative nano-S&T scientists receive considerably more citations both before and after their creative accomplishment, productivity is a poor predictor in this regard. Findings also show that creative scientists link up more otherwise disconnected researchers, and show a broader disciplinary spectrum in their scientific work than their nano-S&T peers. These results suggest that it is not only the sheer quantity of publications that causes scientists to produce creative pieces of work. Rather, their ability to effectively communicate with their colleagues and to address a broad work spectrum are important dimensions in the process of how creative ideas develop.

This paper is part of a larger international research project that aims at understanding the organizational and institutional conditions of creativity in science. The substantial changes seen over the last three decades in the institutional and organizational conditions under which scientific research is conducted give impetus to the desire to know more about the factors that contribute to research creativity. For example, while public research funding was traditionally allocated through long-term institutional block grants to research laboratories and through disciplinary awards to individual academic scientists, lately, competitive project funding has grown considerably. There is also greater emphasis on fostering organized research centers, networks, and interdisciplinary teams.

In an earlier paper, we addressed research creativity by developing a functional typology that brings theoretical, methodological, and empirical aspects of scientific research – each of which has a different function in the research process – into five major categories of creative research accomplishments. This typology was tested in two broad fields of science (HEINZE et al., 2007). Furthermore, we are currently conducting twenty in-depth case studies on organizational and institutional factors that shape effective research environments. The bibliometric analysis of individual nano-S&T scientists presented in this paper is an integral part of a longitudinal multi-method research design that is based on survey, interview, archive, and bibliometric data, and on both quantitative and qualitative research methods, such as network and regression techniques, and in-depth interview analysis.

The next section introduces definitions of creativity and reviews the main theoretical arguments from which these hypotheses are derived. Data and methods for testing these hypotheses are then presented. After a discussion of the empirical results, the concluding section summarizes the findings and discusses the implications and insights gained.

Literature review

Creativity is generally defined as the capability of human beings to do things that are novel, original and valuable (AMABILE, 1996: p. 35, STERNBERG, 2003: p. 89). Creativity is of considerable importance in many areas of society, such as the arts, politics, business, and science. In all these fields of human activity, standards of excellence develop, against which new entities are appraised. In the world of science, such standards are set by scientific disciplines and scientific communities as the main cognitive and social structures for knowledge generation and accreditation (WHITLEY, 2000). However, research judged favorably by peers is not always creative, while creative research is not always initially accepted by peers. There is a tension inherent in the criteria used to judge scientific merit, particularly between plausibility, validation, and originality. Whereas criteria of plausibility and scientific validation encourage conformity, the importance attached to originality encourages dissent, because while scientific originality springs from scientific tradition, it also supersedes it.

Much creativity research has been conducted at the individual level (STERNBERG, 2003; WEINERT, 2000; AMABILE, 1996). Studies that examined the relationship between intelligence and creativity show that while creative people tend to show above average intelligence (as measured by standardized "intelligence quotient" or IQ tests), people with high IQs are not necessarily creative individuals (STERNBERG, 2003). The literature also points to certain behavioral traits that distinguish creative individuals from their peers, such as a high level of curiosity, willingness to learn from experience, preparedness to take risks, persistence in situations of failure, high levels of energy, and distinctive goal-orientation. As both a result and a precondition of these traits, creative people typically tolerate contradictions, ambiguities, and uncertainties in their work (WEINERT, 2000; STERNBERG et al., 1997).

There are two influential sets of arguments in the literature that are particularly helpful starting points for addressing the question of how new ideas and novel science emerge. The first argument is taken from Simonton's chance figuration theory (SIMONTON, 1999; 2004); the second argument from Burt's theory of structural holes (BURT, 1992; 2004).

Simonton argues that highly prolific scientists are more successful in producing high-impact work compared with their less productive peers (SIMONTON, 2004: pp. 14–39). The author offers an intriguingly simple explanation for this fact. "The scientific literature appears to support the conclusion that the quality-quantity relation is best described by the linear function $H = pT$ $(0<p<1)$", where H is the total number of high-impact contributions by a scientist, p the probability of a paper being selected as high-impact, and T the total number of papers published by the scientist (SIMONTON, 2004: p. 23). Simonton's formula is interesting, because the probability p of a single paper to be picked as a creative accomplishment is very low and, in general, follows a Poisson distribution. "By implication, the output of a creative product in a given year must be considered a relatively improbable event for the vast majority of scientists" (SIMONTON, 2004: p. 27). Consequently, if scientists want to increase their number of creative contributions, they need to publish more articles. Simonton concludes that because the low probability p is given, scientists can increase their number of creative and high-impact work only by increasing their publication output.

Simonton's argument is anchored in an evolutionary perspective in which publications are regarded as "ideational variations" of individual scientists who continuously link knowledge elements from their cognitive domain (conceived as a "population of ideas" – phenomena, facts, concepts, variables, constants, techniques, laws, questions, goals, and criteria) into new combinations. The probability p determines the likelihood with which those variations successfully pass several selection filters (e.g. journal peer review) and are retained in the collective stock of knowledge. Since p is miniscule, the number of ideational variations a scientist produces increases her chances that one of the papers is a hit. According to the author, the intertwined relationship between the probability of creative accomplishments in science and individual research productivity pertains to scientific domains as diverse as mathematical logic, physics, biology, psychology, and technology (SIMONTON, 2004: p. 25). Although Simonton acknowledges that what he calls "genius," "logic," and "zeitgeist" are also influential in shaping the emergence of novel science, he subsumes these alternative explanations under the general statement that scientific creativity is a "probabilistic consequence" of research quantity (SIMONTON, 2004: pp. 14–39).

The second influential argument is taken from Burt's theory of structural holes, developed to explain differences in the performance levels of company managers (BURT, 1992; 2004). Burt's argument is also statistical, but from a positional point of view. He argues that individuals who live in the intersection of "social worlds" are more likely to be familiar with selecting and synthesizing cognitive alternatives into "good ideas". Since, according to Burt, thinking and behaving are homogeneous in densely connected groups, people who connect such groups are more likely to be exposed to alternative ways of thinking and behaving, which in turn allows them to make use of varying views, information, and perspectives in their judgments. These people link otherwise disconnected groups and thus bridge what Burt calls "structural holes".

There are several studies that find evidence in favor of Burt's theoretical claims. CROSS & CUMMINGS (2004) demonstrate a positive correlation between performance and "betweenness" among engineers. RODAN & GALUNIC (2004) find that managers' innovation is correlated with the sparseness of their network. Burt himself finds, for a large US electronics company, that "managers whose discussion networks more often spanned structural holes were more likely to express their ideas, less likely to have their ideas dismissed by senior management, and more likely to have their ideas evaluated as valuable" (BURT, 2004: p. 349). Consequently, it is the boundary position of certain individuals which allows them to select and synthesize alternative information and knowledge embedded in internally integrated groups. Individuals who bridge "structural holes" have access to multiple views, information, and perspectives, a fact that explains why they develop more novel and better ideas than their peers. In sum, individuals who occupy a unique position at the nexus of diverse information flows have more opportunities to generate new ideas.

Simonton's and Burt's arguments have not been linked systematically, but it would be highly desirable to know what type of network structure (e.g. brokerage, network size) is positively correlated with research productivity, and which of the two factors more strongly influence the probability of a scientist accomplishing research breakthroughs. The ongoing controversy on the impact of research collaboration on scientific productivity demonstrates, for instance, that it is worthwhile to consider more than one model and one kind of operational structure for the explanation of dependent variables over time. While earlier studies show that frequent collaboration among scientists increases their productivity, LEE & BOZEMAN (2005) find that only the simple number of peer-reviewed journal papers is strongly and significantly associated with the number of collaborators – and thus the size of the co-author network – whereas fractional article count (where co-authors receive the share in the publication count that is equivalent to 1 divided by n authors) is not a significant predictor of publishing productivity.

Hypotheses

For the purpose of this paper, Simonton's and Burt's claims are brought into a set of four hypotheses. The first two hypotheses refer to Simonton's claim of a strong link between creativity and the number of citations and publications. The second set of hypotheses refers to Burt's theory of creativity as an outcome of brokerage in networks. While the first two hypotheses are directly inferred from Simonton's study, the third transfers insights from manager networks to networks of scientists. Burt and his proponents presented empirical evidence on manager networks (see preceding section). Consequently, this is the first time that Burt's theory is put to an empirical test for the world of science. The fourth hypothesis is an analogy in that disciplinary research areas in the sciences are conceived of as densely connected groups. If researchers contribute to several such disciplines, they bridge cognitive boundaries and should be more familiar with intellectual alternatives in their scientific work.

H1: Creative scientists can be predicted from their citation record.

H2: The most important predictor of research creativity is the number of publications.

H3: Compared to their peers, creative scientists link up many more disconnected scientists in research networks.

H4: Compared to their peers, creative scientists show a broader disciplinary spectrum in their scholarly work.

Data and methods

Dependent variable

While previous studies usually relied on a single indicator to identify creative research accomplishments, such as citation and publication data (SIMONTON, 1999; 2004), or prestigious science awards (HOLLINGSWORTH, 2002; 2004), here the dependent variable derives from a combination of survey nominations of highly creative research and scientific prize winners in the field of nano-S&T. Nomination data was collected through an international survey in 2005 where several hundred experts, among them highly cited scientists, active researchers from academia and industry, and editors of major research journals, were asked to nominate creative research accomplishments in their respective fields. A data set of scientific award winners in the field of nano-S&T was then compiled by screening professional societies in Europe and the United States, for instance, the Royal Society, the Royal Swedish Academy of Sciences, the Deutsche Physikalische Gesellschaft, the Société Francaise de Chimie, the American Physical Society. Furthermore, major funding bodies and research organizations were examined, such as the Deutsche Forschungsgemeinschaft, the Max-Planck-Gesellschaft, the Centre National de Recherche Scientifique, the Philip Morris Foundation, and the National Science Foundation (cf. HEINZE et al., 2007). We related these nominations of scientists and groups to the data on prize winners, and thereby derived categories of creative scientists with multiple survey nominations, multiple prize awards, and multiple combinations of survey nominations and prize awards (Table 1).

Table 1. Distribution of creative scientists, combining survey nominations and prize winner data

	Nano-S&T		
	Europe	United States	Total
Multiple prize winners	9	5	14
Multiple nominations	7	21	28
Prize winner and nomination	16	17	33
Multiple prize winners and multiple nominations	3	4	7
Total highly creative scientists	22	29	51
Total scientists in database	224	204	428

Source: CREA data base 2005 (HEINZE et al., 2007)
Note: due to overlap between categories, the total of highly creative scientists is lower than their sum.

There are 51 target scientists in the field of nano-S&T whose research accomplishments took place in the period from the late-1980s until 2004 (Table 1). Those scientists whose creative contributions fall in the periods of 1996–1998 or 1999–2001 were identified, with $N_1 = 33$ scientists in total. Selecting these two time windows was necessary to construct a longitudinal database with at least two observation periods preceding the creative contribution. Since the take-off of the nano-S&T field dates back to the late-1980s, there are scientists with major contributions while the field expanded.

Information on the time period of the creative event was either retrieved from the nomination survey or by detailed analyses of CVs and websites of the respective individuals.

In addition, a comparison group for the N_1 group was constructed, consisting of $N_2 = 33$ peer nano-S&T researchers with the same publishing productivity as N_1 scientists (measured by the number of SCI papers) in the period preceding the creative contribution. Hence, our sample contains matched pairs of equally prolific researchers for whom we collected data in five consecutive three-year periods: 1990–1992, 1993–1995, 1996–1998, 1999–2001, and 2002–2004.

The dependent variable is a dummy with values of "1" for the periods in which N_1 scientists had a research breakthrough (as measured by convergence criterion in Table 1), and values of "0" for the periods in which N_1 scientists had no such creative events. The variable is coded "0" for all N_2 scientists for whom we observe no creative contributions in the periods of 1996–1998 and 1999–2001.

Explanatory variables

All independent variables are taken and constructed from the Science Citation Index (Web of Science Expanded Version) for five consecutive three-year periods: 1990–1992, 1993–1995, 1996–1998, 1999–2001, and 2002–2004. Our variable set includes number of publications, number of citations, degree centrality, an index of network brokerage, and two indices of multidisciplinarity.

Number of publications. The publishing activity of scientists is measured by the number of publications in a given three-year time period. A publication is defined as an article, review, note, or letter. Other publication categories available in the SCI are not considered.

Ln number of citations. Citations are measured by the natural logarithm of the number of citations an author received by December 2005 for the total number of publications he or she published in any of the preceding three-year time periods. This raw number of citations is standardized by a denominator that represents the number of years after their publication. To give an example: an author who published 7 articles in the year 1990–1992, and who received 56 citations for these 7 publications by 2005, gets a citation score of 56/14 = 4 for the 1990–1992 period. This standardization is necessary since articles from earlier periods have a higher probability of being cited than more recent articles.

Degree centrality. We measure the size of our N_1 and N_2 scientist's co-author networks by counting the number of their co-authors in any given three-year time period. All numbers are directly inferred from the publication set of each scientist. The definition of co-author network size reflects a standard measure in social network analysis called "degree centrality" (WASSERMAN & FAUST, 1994).

Network brokerage index. The brokerage index measures the percentage of those peer scientists who are unconnected in the nano-S&T publication sample unless connected as co-authors of N_1 and N_2 scientists. If there is only one publication, all authors and co-authors are directly connected and there is no opportunity for brokerage. Therefore, in the first descriptive part of the result section below, only researchers with at least two publications in any three-year period are considered. The network brokerage index can be stated mathematically as below, where n is the number of ego's peer scientists (= alteri), (n^2-n) the number of pairs between alteri without ego (= potential ties), and x_{ij} the number of alteri ties without ego:

$$\text{Network brokerage index} = \frac{(n^2 - n) - \sum_{i=1}^{n} \sum_{j=1}^{n} x_{ij}}{(n^2 - n)}$$

This measure has a theoretical range from 0.0 to 1.0, but ranges in this sample from 0.0 to 0.5. This means that all N_1 and N_2 scientists of this sample ("egos") connect at a maximum rate of 50 percent of those peers with whom they publish in a given time period ("alteris"). In contrast to Burt, undirected data is processed. Therefore, structural holes cannot be calculated directly. Index calculation is based on the normalized broker measure available in the in-neighborhood routine of ego network's density in UCINET 6.0 (BORGATTI et al., 2002). Whether N_1 and N_2 scientists link up otherwise disconnected groups or clusters is not tested. Instead, this measure refers to the inter-individual level of brokerage.

Multidisciplinarity indices. Two different indices of multidisciplinarity are constructed. These indices measure both variety and concentration of N_1 and N_2 scientists' publishing behavior across either SCI subject codes, or SCI journals. Examples for subject codes are applied physics, polymer science, material science, or optics. Examples for journals are *Applied Physics Letters*, *Surface Science*, *Nano Letters*, or *Langmuir*.[1] The index combines the number of subject codes (or journals) in which scientists publish and the concentration of their publications across these subject codes (or journals) using the Gini coefficient. The index increases when scientists publish in different subject codes (or journals); it decreases when they publish most of their work in few subject codes (or journals). Consequently, if scientists publish across various subject codes but show a high concentration in a few, they receive lower values than those scientists with a more equal distribution in their publishing activity.

[1] Full lists of SCI subject codes (N=170) and SCI journals (N= ca. 2500) used for this index can be requested from the first author.

This can be formalized into the following formula, where x_i denotes the number of subject codes (or journals):

$$\text{Multidisciplinarity index} = (1 - GINI)\sum_{i=1}^{n} x_i \cdot$$

For instance, if a scientist had 10 publications evenly distributed in 5 subject codes in one time period, our index gives a score of (1-0)*5 = 5, because the Gini coefficient is 0. However, if 6 of the 10 papers are published in one subject code, while the four others are distributed across four other subject codes, the Gini coefficient is 0.4 which gives an index score of (1–0.4)*5 = 3.

Results

Descriptive longitudinal statistics

The descriptive findings give considerable support to the four hypotheses stated above. Over a period of fifteen years, N_1 scientists publish distinctly more articles than their N_2 peers (H2), even though both groups are matched on the number of publications in the period preceding the creative event. There is also a stronger citation record for N_1 scientists (H1), and conspicuously higher values in the brokerage (H3) and multidisciplinarity indices (H4). The inter-group comparison of N_1 and N_2 scientists offers the first empirical evidence on the influence of these independent variables on research creativity.

The matching of the two groups of N_1 and N_2 scientists is clearly depicted in Figure 1, showing identical values for the period preceding the creative event (–1). Both groups publish, on average (median), 4 papers. In the periods following the creative event (CE, +1, +2), however, N_1 scientists publish, on average (median), 12 papers more than N_2 scientists. While the productivity of N_1 scientists increases substantially after the creative event (CE), N_2 scientists do not increase their publishing productivity. In the three latter periods (CE, +1, +2), all T-tests on mean differences between creative scientists (N_1) and the matched group (N_2) are statistically significant on the 0.05 level. The fact that N_2 scientists publish less after period CE might be explained by their move to industry or to non-research jobs. If the selection mechanism in science works effectively, such job moves would support our hypotheses. We are currently collecting CVs of N_1 and N_2 scientists to examine this issue.

Both groups also differ considerably in their citation scores. The citation scores of N_1 scientists are much higher than those of the N_2 scientists, not only in the two preceding periods but also in the two periods following the CE. While N_1 scientists receive 39 standardized citations before and 145 standardized citations after the CE, the values for N_2 scientists are 12 and 21 respectively. T-tests on the mean differences of citation scores between N_1 and N_2 are all significant on the 0.05 level.

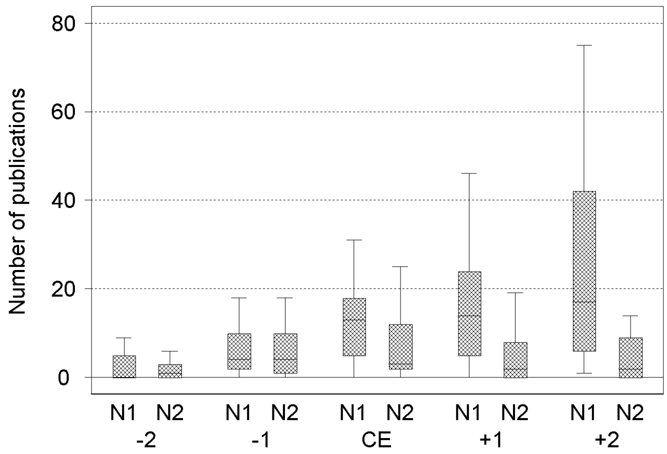

Figure 1. Longitudinal publication scores of N_1 and N_2 scientists
Note: All authors, excluding outliers

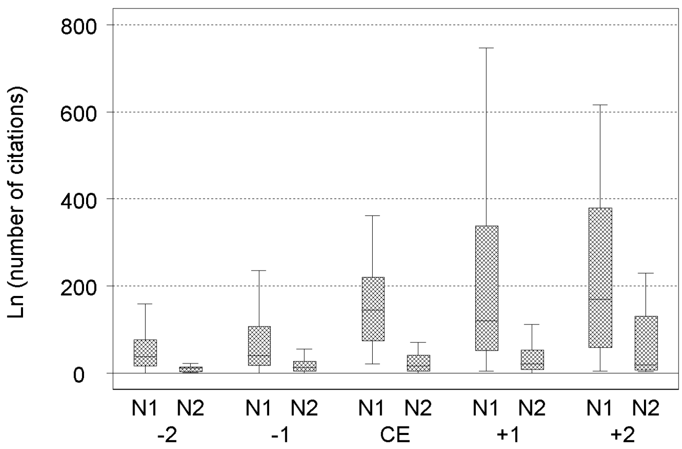

Figure 2. Longitudinal citation scores of N_1 and N_2 scientists
Note: authors with at least 2 publications, excluding outliers

Similarly, the size of the co-author network clearly differs between the two samples (Figure 3). Scientists who have made highly creative contributions to science tend to have larger co-author networks than their equally productive matched peers.

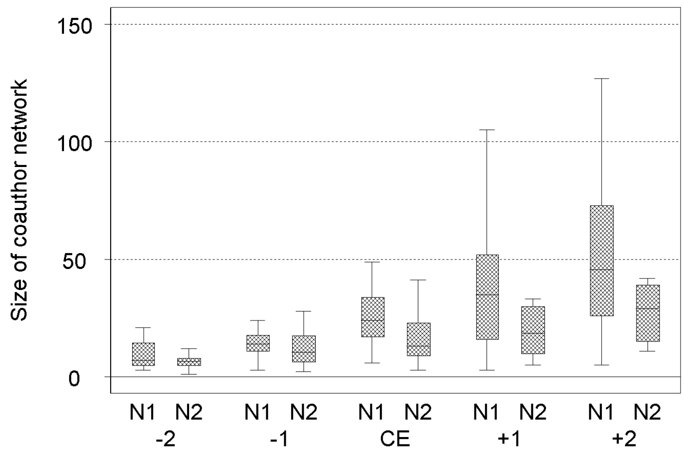

Figure 3. Size of co-author network of N_1 and N_2 scientists
Note: authors with at least 2 publications, excluding outliers

In the period of the CE, creative scientists belong to a network of about 26 colleagues, on average (mean). N_2 researchers, however, have an average network size of only 17 co-authors. These differences are statistically significant on the 0.05 level. Evidence (below) suggests that degree centrality has only little explanatory power in the dependent variable when other key variables are introduced in the regression model.

Figure 4 shows the longitudinal development of the normalized brokerage index scores for both N_1 and N_2 scientists. N_1 scientists have, on average (median), higher brokerage values and smaller ranges than N_2 scientists in all five periods. Note that the variance of N_2 scientists is much higher, and that both distributions overlap considerably. Therefore, T-tests on the mean differences between N_1 and N_2 are significant on the 0.05 level only in the –2 and CE period. In sum, there is limited evidence that N_1 scientists are more active brokers in their publication networks.

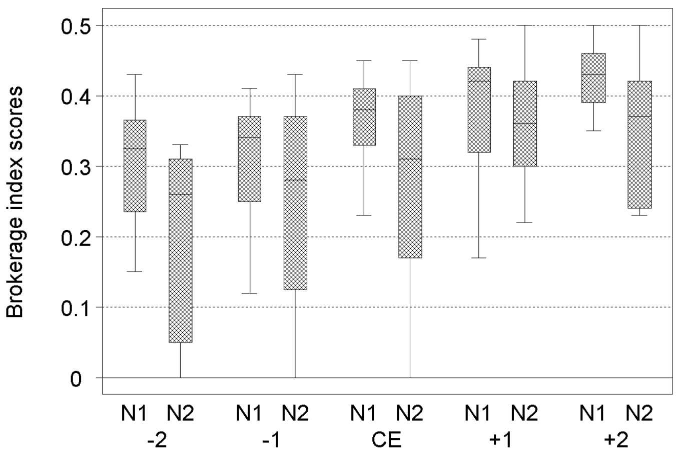

Figure 4. Brokerage index scores of N_1 and N_2 scientists
Note: authors with at least 2 publications, excluding outliers

There are further differences between N_1 and N_2 scientists with respect to the two multidisciplinary indices (Figures 5 and 6). When using SCI journals (Figure 5), N_1 scientists have a score of 5.0 in the periods preceding the CE (–2, –1, CE), while N_2 scientists' score is about 3.1. In the periods following the CE (CE, +1, +2), N_1 scientists have a score of 7.9 but N_2 scientists score only 4.1. T-tests on the mean differences of the brokerage index between N_1 and N_2 are significant in the –1, CE and +2 periods on the 0.1 and 0.01 levels. However, there is less difference between N_1 and N_2 scientists with respect to the subject code index where distributions for both groups overlap substantially (Figure 6). Hence, mean differences are statistically significant on the 0.05 level only in –2 and CE periods.[2]

In sum, there is some evidence that scientists whose work spans a wide range of academic journals are aware of a richer set of information and perspectives that, in turn, enables them to publish results which are valued as creative by their peer scientists (as visible in nomination and prize data). These scientists are capable of speaking to different audiences and specialties, so their work can be used more widely than that of more specialized scientists. This conclusion is further supported by the citation data indicating above average recognition of N_1 scientists (Figure 2).

[2] Note that values of this variable are smaller due to a lower number of SCI subject codes compared to SCI journals.

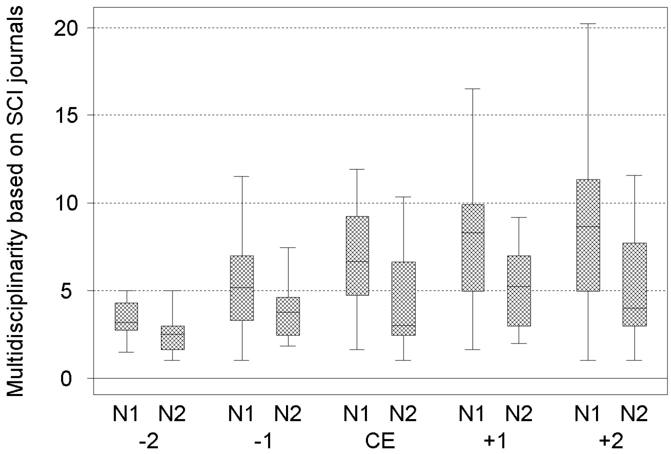

Figure 5. Multidisciplinarity scores of N_1 and N_2 scientists (SCI Journals)
Note: authors with at least 2 publications, excluding outliers

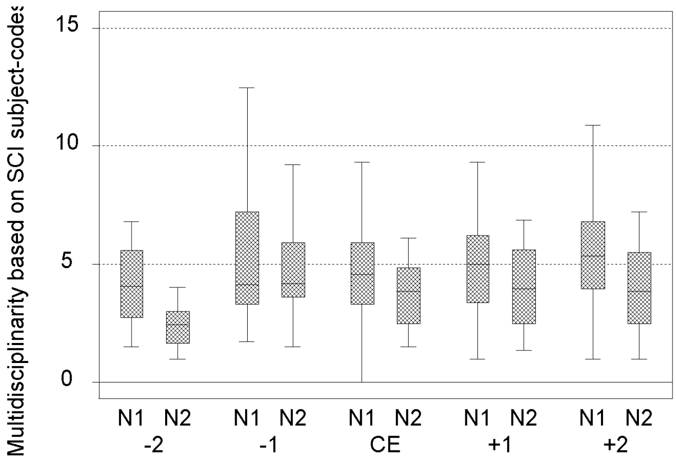

Figure 6. Multidisciplinarity scores of N_1 and N_2 scientists (SCI Subject Codes)
Note: authors with at least 2 publications, excluding outliers

Longitudinal regression analyses

Pooled regression model (inter-personal differences). The descriptive longitudinal findings are summarized in a pooled regression model, estimated in STATA 9.0 (Table 2). Except for the number of publications, all variables exert a significant influence on research creativity. Citations, network size, and multidisciplinarity (journal index) increase the probability of a given scientist conducting and publishing scientific work that is judged creative by his peers. In contrast, the number of publications, the network brokerage index, and the second multidisciplinarity measure (subject code index) exert a negative influence (a coefficient below 1). These results clearly confirm H1, but they do not support H2. So, evidence for Simonton's claims is mixed. There is mixed evidence also for H4, while H3 is apparently rejected.

Table 2. Pooled logit regression on creativity (odds ratios)

Publications	1.13***	1.00	0.95	0.96	0.98	0.91
Ln(citations)		2.23***	2.04***	2.52***	2.67***	2.77***
Degree centrality			1.06***	1.09***	1.11***	1.09***
Network brokerage				0.00***	0.01**	0.00***
Multidisciplinarity (subject codes)					0.77***	0.59***
Multidisciplinarity (journals)						1.67***
Constant	0.07***	0.01***	0.01***	0.01***	0.02***	0.02***
Pseudo R^2	0.15	0.27	0.30	0.33	0.35	0.38
N	210	210	210	210	210	210

* p<0.1; ** p<0.05; *** p<0.01

Fixed effects regression models (intra-personal differences). Characterizing and comparing N_1, with N_2 scientists does not elucidate the causal mechanism that makes N_1 scientists creative over time. In contrast to the pooled models presented above, longitudinal intra-personal comparisons are able to determine those explanatory variables which are strong enough to cause substantial change in the variation of the dependent variable of a given individual over time. For this reason, fixed effects regression models are estimated in STATA 9.0 that determine those independent variables that cause the dependent variable to switch from 0 to 1 within the group of N_1 scientists. Fixed effects regression models "time-demean" the data, i.e. variables are transformed by subtracting the mean from values, so only the within-variation is left (ALLISON & WATERMAN, 2002). Since N_2 scientists have no variation in the dependent variable over time, they are dropped from further consideration. These models are longitudinal, examining N_1 = 33 scientists over 15 years with five consecutive three-year time periods: N = 165.

First, a conditional LOGIT regression model analyzes the extent to which the number of citations (H1), the number of publications (H2), network brokerage (H3), and multidisciplinarity (H4) explain a scientist's probability of accomplishing creative work (Table 3). Since the time period in which the CE occurred is known, the influence of explanatory variables in preceding periods can be tested. The size of the co-author network (degree centrality) and number of publications are control variables.

The results of the conditional LOGIT model show that the regression coefficient (odds ratio) for the citation score is highly significant and explains considerable variations in the dependent variable. Multidisciplinarity, measured by journals, also increases the likelihood of scientists to accomplish creative work, but the coefficient is only weakly significant. In contrast, while multidisciplinarity (subject code index), network brokerage, and number of publications have negative influences on the dependent variable (values below 1), their influence is not statistically significant throughout the model steps. In sum, the fixed effects model suggests that citations are the most important mechanism in explaining why certain scientists produce novel ideas in science.

Table 3. Conditional logit regression on creativity (odds ratios)

Publications	1.01	0.94***	0.99	0.99	0.99	0.94
Ln(citations)		3.19***	3.13***	3.45***	3.45***	3.34***
Degree centrality			0.97	0.97	0.97	0.98
Network brokerage				0.17	0.18	0.04
Multidisciplinarity (subject codes)					0.99	0.84
Multidisciplinarity (journals)						1.33*
N	165	165	165	165	165	165

* p<0.1; ** p<0.05; *** p<0.01

Table 4. OLS regression on ln(citations) with fixed effects

Publications	0.08***	0.06***	0.05***	0.05***	0.04***	0.04***	0.04***
Degree centrality		0.01	–0.01	–0.01	–0.01	–0.01	–0.01
Network brokerage			8.10***	8.07***	7.59***	7.77***	5.87***
Multidisciplinarity (subject codes)					0.00	–0.03	0.04
Multidisciplinarity (journals)						0.05	0.01
Episode –1							0.11
Episode CE							1.14***
Episode +1							0.93***
Episode +2							0.54*
Constant	2.40***	1.65***	1.64***	0.70***	0.68***	0.61***	0.61***
R^2 (within)	0.42	0.42	0.74	0.74	0.74	0.74	0.79
N	165	165	165	165	165	165	165

* p<0.1; ** p<0.05; *** p<0.01

Further analysis of the citation variable reveals that network brokerage is an important mechanism for research creativity, but that it works indirectly. A fixed effects OLS regression model can be calculated with standardized citation scores as the dependent variable, using the remaining independent variables as explanatory factors (Table 3). Results are straightforward. The number of citations is positively influenced both by the number of publications and by the level of network brokerage. The regression coefficients for these two variables are highly significant and explain considerable variation in the dependent variable. Note that the broker variable is much stronger and increases the explained variance of the model considerably (R^2). All other independent variables, except for degree centrality, have positive but insignificant coefficients.[3]

Summary

Comparing longitudinal data for highly creative scientists and equally productive researchers, it is confirmed that creative scientists can be predicted on the basis of the total number of citations (H1). However, Simonton's conclusions have to be specified. First, his statement that the single most critical predictor of high-impact work is the total number of publications (H2) is too simple, because the broker effect is much stronger in the final OLS model. Scientists who effectively broker otherwise disconnected colleagues receive higher citation scores (H3). Second, being creative in research also depends on the disciplinary scope of individual scientists, at least when measured by our multidisciplinarity index based on journals (H4). In sum, Burt's insights on relational network position add to our understanding of why certain individuals have more creative ideas than others. These findings are synthesized in Figure 7.

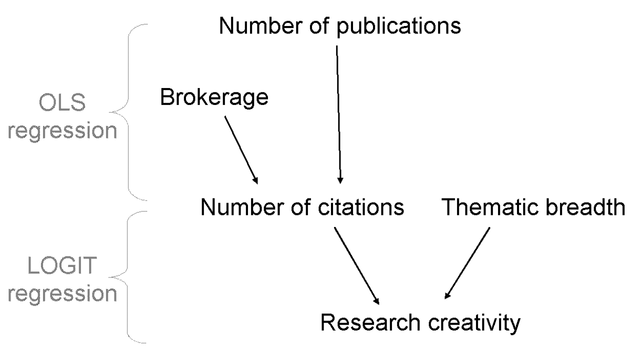

Figure 7. Results from LOGIT and OLS regression analyses

[3] In addition, we calculated the OLS model with both N_1 and N_2 scientists (not documented in this paper). The latter group can be included in the equation since their citation values vary over time. The results are almost identical.

Discussion

This work contributes to the controversy about key factors influencing the research creativity of individual scientists. It is not only the sheer quantity of publications that causes scientists to produce creative pieces of work. Rather, the ability of scientists to effectively communicate with their colleagues and to address a broader work spectrum are important dimensions in the process of how creative ideas develop. These results suggest that there are several predictors for creative science accomplishments. It is worthwhile to triangulate hypotheses derived from Simonton's and Burt's theories, and to put them to an empirical, longitudinal test.

The paper has a number of methodical strengths that validate the results. First, highly creative research contributions are determined independent both of standard bibliometrics and of the explanatory variables using results from an international survey and from science awards. Second, highly creative scientists are compared with a control group matched on the theoretically challenging publication count variable. Third, inter-personal comparisons between N_1 and N_2 scientists, based on descriptive statistics and regression models, are complemented by fixed effects models that shed light, within the group of N_1 scientists, on intra-personal factors that cause the creativity variable to switch from 0 to 1. Fourth, the longitudinal research design covering five consecutive three-year time periods allows the determination of quasi-causal effects.

This paper also has limitations that need to be taken into account when making conclusions and that indicate the need for future research. For instance, a criticism could be that while referring to Burt's theory of structural holes, the broker index used here is not identical with Burt's. While this is true, bear in mind that co-author relations are non-directional ties, so that structural holes measures cannot be computed with this data.

Furthermore, the adequacy of the two multidisciplinarity indices may be questioned. Admittedly, no simple measure for the concept of multidisciplinarity is available as yet. However, based on experience with the subject codes variable, which has much less descriptive and explanatory power than the journal-based variable, the latter is a good approximation. In contrast, there should be some caution that sophisticated delineations of science sub-fields or thematic sub-areas (and nano-S&T spans a wide range of such fields and areas) could stand up to the challenge of adequate operational implementation and measurement in this regard. Nevertheless, further suggestions to deal with the methodical problem of proper field boundaries that go beyond the subject code delineation provided by ISI are welcome.

Perhaps most importantly, the paper addresses scientific creativity on the individual level, while research today is conducted in groups and institutions. Among the factors most commonly believed to be conducive to research environments are: autonomy for researchers, adequate facilities and funding, a variety of disciplines and fields, a wellmanaged staff selection, a flat and decentralized organizational structure, and visionary leadership. Common negative factors include: insufficient basic funding, limited time for research, bureaucratic management, a narrow range of disciplinary expertise, and excessive evaluation and accountability pressures (HEMLIN et al., 2004: pp. 16–17, 195–196). As stated above, we are conducting twenty in-depth case studies on those organizational and institutional factors that shape effective research environments. Results from these qualitative case studies, many of which are in the field of nano-S&T, will complement and add to the insights that we derive from this paper.

*

This paper is based on research undertaken by the Project on Creativity Capabilities and the Promotion of Highly Innovative Research in Europe and the United States (CREA), funded by the European Union *New Emerging Science and Technologies* (NEST) program. The authors gratefully acknowledge contributions to this project by Jacqueline Senker, Cheryl Leggon, Juan Rogers, Philip Shapira and Stefan Kuhlmann. We are also grateful for comments and suggestions from Stephan Gauch and other participants of the *9th Science & Technology Indicators Conference*, Leiden, 2006, and Philippe Laredo and other participants of the *Atlanta Conference on Science and Technology Policy*, Atlanta, 2006.

References

ALLISON, P. D., WATERMAN, R. P. (2002), Fixed-effects negative binominal regression models, *Sociological Methodology*, 32: 247–265.

AMABILE, T. M. (1996), *Creativity in Context*, Westview Press, Boulder, Colorado.

BORGATTI, S., EVERETT, M., FREEMAN, L. (2002), *Ucinet 6 for Windows. Software for Social Network Analysis*, Analytic Technologies, Natick.

BURT, R. S. (1992), *Structural Holes. The Social Structure of Competition*, Harvard University Press, London.

BURT, R. S. (2004), Structural holes and good ideas, *American Journal of Sociology*, 110 : 349–399.

CROSS, R., CUMMINGS, J. N. (2004), Tie and network correlates of individual performance in knowledge-intensive work, *Academy of Management Journal*, 47 : 928–937.

HEINZE, T. (2006), *Die Kopplung von Wissenschaft und Wirtschaft. Das Beispiel der Nanotechnologie*, Campus, Frankfurt/New York.

HEINZE, T., SHAPIRA, P., SENKER, J., KUHLMANN, S. (2007), Identifying creative research accomplishments: methodology and results for nanotechnology and human genetics, *Scientometrics*, 70 : 125–152.

HEMLIN, S., ALLWOOD, C. M., MARTIN, B. R. (Eds) (2004), *Creative Knowledge Environments. The Influences on Creativity in Research and Innovation*, Cheltenham: Edward Elgar.

HOLLINGSWORTH, R. (2002), *Research Organizations and Major Discoveries in Twenthieth-Century Science: A Case of Excellence in Biomedical Research.*, WZB Discussion Paper P02-003, Berlin.

HOLLINGSWORTH, R. (2004), Institutionalizing excellence in biomedical research: The case of Rockefeller University. In: D. H. STAPLETON (Ed.), *Creating a Tradition of Biomedical Research. Contributions to the History of the Rockefeller University*. Rockefeller University Press, pp. 17–63.

HULLMANN, A., MEYER, M. (2003), Publications and patents in nanotechnology. An overview of previous studies and the state of the art, *Scientometrics*, 58 : 507–527.

LEE, S., BOZEMAN, B. (2005), The impact of research collaboration on scientific productivity, *Social Studies of Science*, 35 : 673–702.

RODAN, S., GALUNIC, C. (2004), More than network structure: How knowledge heterogeneity influences managerial performance and innovativeness, *Strategic Management Journal*, 25 : 541–562.

SIMONTON, D. K. (1999), *Origins of Genius: Darwinian Perspectives on Creativity*, Oxford University Press, New York.

SIMONTON, D. K. (2004), *Creativity in Science: Chance, Logic, Genius, and Zeitgeist*, Cambridge University Press, Cambridge.

STERNBERG, R. J. (2003), *Wisdom, Intelligence, and Creativity Synthesized*, Cambridge University Press, Cambridge.

STERNBERG, R. J., OHARA, L. A., LUBART, T. I. (1997), Creativity as investment, *California Management Review*, 40 : 8–32.

WEINERT, F. E. (2000), Individuelle Kreativität und kollektives Ergebnis, *Der Architekt. Zeitschrift des Bundes deutscher Architekten*, 48 : 24–31.

WHITLEY, R. (2000), *The Intellectual and Social Organization of the Sciences, 2nd edition*, Oxford University Press, Oxford.