

Article

An Empirical Test of the Inter-Relationships between Various Bibliometric Creative Scholarship Indicators

Boris Forthmann ^{1,*}  and Mark A. Runco ² ¹ Institute of Psychology in Education, University of Münster, 48149 Münster, Germany² Innovation and Leadership Faculty, Southern Oregon University, Medford, OR 97501, USA; runcoM@sou.edu

* Correspondence: boris.forthmann@wwu.de

Received: 6 April 2020; Accepted: 17 June 2020; Published: 19 June 2020



Abstract: Quantifying the creative quality of scholarly work is a difficult challenge, and, unsurprisingly, empirical research in this area is scarce. This investigation builds on the theoretical distinction between impact (e.g., citation counts) and creative quality (e.g., originality) and extends recent work on using objective measures to assess the originality of scientific publications. Following extensive evidence from creativity research and theoretical deliberations, we operationalized multiple indicators of openness and idea density for bibliometric research. Results showed that in two large bibliometric datasets (creativity research: $N = 1643$; bibliometrics dataset: $N = 2986$) correlations between impact and the various indicators for openness, idea density, and originality were negligible to small; this finding supports the discriminant validity of the new creative scholarship indicators. The convergent validity of these indicators was not as clear, but correlations were comparable to previous research on bibliometric originality. Next, we explored the nomological net of various operationalizations of openness and idea density by means of exploratory graph analysis. The openness indicators of variety (based on cited journals and cited first authors) were found to be made up of strongly connected nodes in a separate cluster; the idea density indicators (those based on abstracts or titles of scientific work) also formed a separate cluster. Based on these findings, we discuss the problems arising from the potential methodological overlap among indicators and we offer future directions for bibliometric explorations of the creative quality of scientific publications.

Keywords: creativity; openness; diversity; idea density; impact

1. Introduction

Findings from empirical research are reliable because the scientific method includes a variety of procedures for quality control, such as peer review [1,2] and replication [3,4]. Further, important findings are still vetted after publication, as scientists read published research, whereby they cite some findings and ignore others. This leads to some findings that have more “impact” than others [5], as highlighted in Garfield’s groundbreaking work [6,7]. Since then, impact has gained incredible popularity: it is now used to quantify scholarship, in that it is used in ratings of individual scientists [8,9] and to compare journals [10]. It is even used in some tenure and promotion decisions [11] and taken into account by many grantors (for critique of this practice, see van den Besselaar et al. [12]).

Impact may be used so often in part because it can be studied objectively (e.g., citation counts [13]), but impact is distinct from creativity [14–18]. While impact is related to reputation and fame and other influencing factors [19], creativity is related to originality. For example, if one aims to assess the creativity of a product, one might typically examine its originality and effectiveness [20]. This distinction between impact and creativity in the context of scientific work has been examined by Heinze [21], who used the terms *originality* and *scientific relevance* (i.e., effectiveness in the context of science) to

distinguish research accomplishments according to the following categories: ignored (low-originality and low-effectiveness), mainstream (low-originality and high-effectiveness), contested (high-originality and low-effectiveness), and creative (high-originality and high-effectiveness). Here, we aim to extend this work by investigating the traits of scientific research that are related to creativity. Namely, similar to how the psychological research on creativity has identified various key traits related to the creativity of individuals, including openness, risk tolerance, autonomy, and intrinsic motivation [22,23], in this work we attempt to study analogous bibliometric indicators of creative scholarship that may, in turn, be useful to highlight the key facets of creativity relevant at the level of individual scholars [9].

Because some indicators of creativity can be estimated from bibliometric data, they can be used to bring research on scholarship closer to creativity (i.e., these indicators can be used to incorporate the concept of creativity into research on scholarship), which may minimize the conflation of impact and creativity [17]. In the present research, we identify several indicators that can be objectively determined and that are more aligned with creativity than impact, and we examine these indicators and their relationship with impact and originality.

1.1. Quantifying Aspects of Creativity in Science

In the fields of both creativity research [24] and bibliometrics, investigations have been done that used citations as the data source. Simonton [25,26], for instance, reported that the number of publications for any one individual is correlated with the number of citations received by that individual [27–29]. He suggested that the number of citations could thus be used as a measure of research quality and that this could be predicted from the productivity measures (i.e., quantity). Simonton concluded that bibliometric research probably only needs to examine quantities, because they are so highly correlated with measures of quality. A largely similar view on citations was offered by Wang [30]. Of note though, in these studies the quality measure is synonymous with impact, such that this idea of quality is not really a measure of originality or creativity.

Nonetheless, scientometric studies have acknowledged the distinction between impact and aspects of creative quality, such as originality, significance, depth, correctness, completeness, and clarity [31,32]. Long [33] examined comparable criteria related to creative quality as applied by judges with varying degrees of expertise (undergraduates to researchers) to assess how people responded to tasks designed to measure scientific thinking. Furthermore, Azoulay et al. [34] examined the keywords of published research and suggested that newness is indicative of originality; they interpreted any new keywords introduced in a particular article as being indicative of an original piece of research. A parallel approach was used in another study, which examined the information contained in abstracts within certain disciplines and used any newly introduced concept, method, or element as an index of quality [35]. Furthermore, in a study by van den Besselaar et al. [12], who developed an independence index at the author level to reflect the quality of early career researchers based on their degree of scholarly independence (i.e., having their own co-author network, level of thematic independence, and focus on new topics), novelty was assessed by the growth of the individual's research topics and fields (i.e., fast growth of both indicates promising research [12]); this is strongly connected to the use of novel keywords [34,36]. Finally, Funk and Owens-Smith [37] examined inventions, including ones that consolidated existing streams of knowledge and others that disrupted existing streams (also see Wu et al. [38]).

One study in particular used bibliometric data to assess originality. Shibayama and Wang [39] developed a method to estimate originality by comparing the references of a target publication to the subsequent references to that target publication. The reasoning here was that if a publication was original, it would contain some knowledge that was not available previously. Shibayama and Wang tested this using a set of publications and then validated the resulting index by examining its relationship with questionnaire data from 268 PhDs in the life sciences [39]. These PhDs were asked to rate their own dissertations on both theoretical originality and methodological originality, where each was measured on a three-point scale. These ratings were correlated with the originality scores from the

new bibliometric index, which was based on articles the PhDs had published (as the first or second author) one to two years before they completed their doctorates. Shibayama and Wang compared three indices of originality: one was uncorrected, as described above, but the second and third were corrected to account for (a) the number of references in the research citing the target paper or (b) the number of citations received by the references in the target paper, as well as the number of references in the research citing the target paper.

Analyses indicated that the Shibayama and Wang indices significantly correlated with the self-ratings of theoretical originality by the PhDs (most of the correlations were approximately in the range from 0.10 to 0.20) but not with methodological originality. Importantly, subsequent regression analyses demonstrated that originality was correlated with self-assessed creativity even after the number of citations and references were taken into account. Thus, even though this approach focused on newness, which is a synonym of originality, and it was objective, concerns remain. First, the primary validation of the originality indices was a self-assessment of originality, and self-assessments have various biases (e.g., memory, honesty, social-desirable responding; see Anastasi [40]). Next, potential problems may arise from the use of references and citations. Shibayama and Wang were well aware of these issues, as they acknowledged, for example, that an article may be cited for refutation (because it made a mistake), not because it presents new valuable knowledge [39], or an article may be cited for non-academic reasons (i.e., “social and political motivations,” p. 5) rather than because it introduces new information. Further, some disciplines may have the tradition of citing earlier work either more liberally or more stringently.

1.2. Aim of the Current Work

The present research is designed to examine creativity rather than impact, and for this reason we draw heavily on the creativity research [22–24]. This research suggests that various indicators available in bibliometric data are strongly related to creativity. Thus, one of our aims is to test the possibility of using these new indicators. Another aim is to see how the new indicators relate to those proposed by Shibayama and Wang [39].

Creative individuals and the creative process are often categorized by “openness” [41–43]. We reason that a scientist who has a tendency toward openness will consider various perspectives and cite a wide range of authors and journals; thus, we take the diversity of research cited by an author as an indicator of openness. Specifically, we use two sources of information from an article’s references to quantify openness: (a) the journals from which research was cited and (b) the first authors whose research was cited. This information is quantified using common diversity indices in the bibliometric literature, namely variety, disparity, and the Rao-Stirling index [44,45].

Both empirical research and theory also suggest that ideation plays a role in the creative process. Various tests of ideation [46,47] exist, but one can also calculate the idea density of a written work. This calculation, which originated in linguistic studies, was later used by Runco et al. [48] in a series of studies on creativity. The present investigation uses two estimates of idea density.

In sum, the main aim of this work is to determine whether a set of objective indices of creative scholarship could be reliably extracted from Web of Science (WoS) data and to determine whether these creative scholarship indicators could be correlated with impact and with a previously published measure of bibliometric originality. Recall here that originality is a necessary but insufficient aspect of creativity [20,21], and that creativity is distinct from impact [14–18]. These conclusions from previous research lead us to expect that the new creative scholarship indicators proposed here will be positively correlated with originality but will be much less strongly related to impact.

2. Materials and Methods

2.1. Datasets

2.1.1. Creativity Research Articles

The first dataset was retrieved from the four major journals in creativity research (arguably representative of work from this field): the *Journal of Creative Behavior*, the *Creativity Research Journal*, *Thinking Skills and Creativity*, and *Psychology of Aesthetics, Creativity, and the Arts*. These four journals were chosen because this was an initial analysis using the outlined approaches (see below) and one cluster of journals which all focus on one subject matter was a reasonable starting point. Given that the examined variables have implications for creative scholarship, creativity was chosen as the subject matter, and these four journals are central to that field. In addition, ten-year periods are commonly used in studies based on publication and citation-based indicators [19] and, hence, only articles (i.e., no other document types) from the period from 2009 to 2019 were retrieved from the WoS (<https://apps.webofknowledge.com>). The following search was used:

PUBLICATION NAME: (psychology of aesthetics creativity and the arts) *OR* **PUBLICATION NAME:** (thinking skills and creativity) *OR* **PUBLICATION NAME:** (creativity research journal) *OR* **PUBLICATION NAME:** (journal of creative behavior)

Refined by: **PUBLICATION YEARS:** (PY = 2009–2019) *AND* **DOCUMENT TYPES:** (ARTICLE).

This resulted in a total of $N = 1643$ articles.

2.1.2. An Update of the *biblio* Dataset

The *biblio* dataset is included in the R package *bibliometrix* [49], but for the purpose of the current research, a larger dataset was required (the *biblio* dataset only includes 99 articles). Hence, the following search string was used to enlarge the search scope:

TITLE: (bibliometric*)

Refined by: **PUBLICATION YEARS:** (2007 *OR* 2019 *OR* 2006 *OR* 2018 *OR* 2017 *OR* 2016 *OR* 2015 *OR* 2014 *OR* 2013 *OR* 2012 *OR* 2011 *OR* 2010 *OR* 2009 *OR* 2008) *AND* **DOCUMENT TYPES:** (ARTICLE)

Analogous to the creation of the *biblio* dataset, this search focused on titles; but, here “bibliometric” was combined with an asterisk to include all words with the stem *bibliometric*. Again, articles were only retrieved from the time span between 2006 and 2009. This resulted in $N = 2986$ documents for replicating the findings obtained from creativity research.

2.2. Measures

2.2.1. Impact

The number of citations that a publication received as documented in WoS was used as an indicator of impact. More recent publications are less likely to receive citations [50] and, hence, we regressed citation counts on year of publication and used the resulting residuals in the analysis. Hence, we fitted models with polynomial terms [50] for year of publication to predict citation counts and chose the residuals from the best fitting model. This was a cubic model for the creativity dataset and a quadratic model for the bibliometric dataset.

2.2.2. Openness

Empirical research and theory posit that openness plays a key role in the creative process [42,51]. Here, as also justified by empirical findings and creativity theory [52,53], we focused on an author’s openness to diverse sources of information. Importantly, “openness”, as it used in this work, should not be confused with terms such as “open science” or “open access”. Operationally, the indicators used here were derived as diversity measures and were based on two different sources of information,

both taken from the references. One, an author's openness to diversity, was based on the number of different journals cited in the references. Second, an author's openness to other researchers' works, was based on the diversity of cited first authors in the references. Given that many formulas exist to calculate diversity [44,45], we chose ones with the closest conceptual connection with creativity. First, we calculated the Rao-Stirling index based on different journals (RS-JN) and different first authors in the references (RS-FAN). The Rao-Stirling index is a composite index that accounts for all commonly used factors to quantify diversity, namely variety, balance, and disparity [44]. Of these factors, variety and disparity seem to be more closely connected with creativity. That is, based on the associative theory of creativity [54], creative work should emerge from combining more distant concepts (analogous to ideation research; e.g., Beketayev and Runco [55]). Arguably, combining ideas from different journals and first authors are promising sources of information from this perspective. Hence, we calculated the variety and disparity for both different journals (V-JN and D-JN) and first authors (V-FAN and D-FAN) in the references. All these indices were calculated by means of the R package *diverse* [56]. That is, variety was the number of category counts (i.e., the number of different cited journals and the number of different cited first authors).

Disparity for different cited journals and first authors was based on Schubert's [57] Hirschian similarity measure and Jaccard similarity based on an author coupling network matrix, as implemented in the *biometrix* R package [49], respectively. The Hirschian similarity between journals requires downloading of the Citing Journal Data from Journal Citation Reports (JCR; <https://jcr.clarivate.com>), which include a list of journals cited by a given journal. The *h*-cores from these lists of any two pair of journals are then used to derive a Jaccard-type similarity measure between journals [57]. First, the creativity dataset was scored. A total of 6201 sources cited by the articles included in the creativity dataset were screened for availability of Citing Journal Data in JCR. From this list, 2129 Citing Journal Data were used to create a journal similarity matrix. Then, disparity and Rao-Stirling index based on these similarities were calculated for all available journal data and for journals that were cited more than 10 times. The correlation between disparity based on the full data and disparity based on the journals that were cited more than 10 times was $r = 0.94$, and the correlation between both Rao-Stirling indices (i.e., based on full and reduced data, respectively) was $r = 0.99$. Based on these correlational findings it was decided to retrieve Citing Journal Data for the bibliometric dataset only for journals that were cited more than 10 times. Hence, for the bibliometric dataset a total of 1451 sources were screened for availability and 1235 were retrievable for analysis. Journal similarity for both datasets was based on available data, with the only exception of similarity between exactly matching sources, which was always set to a value of one. Moreover, similarity of first authors was calculated only for first authors, who were also among the authors of the articles included in both datasets (i.e., active authors in the fields of creativity or bibliometric research). For this similarity measure a Jaccard-type similarity measure based on the author coupling matrix was also used [49]. Analogous to the journal similarity measure, exactly matching author names were also set to a similarity value of one (i.e., also for authors who were not active authors in the field).

Similarities based on journals and first authors were then transformed into disparities by 1-similarity [58] and in this study we used the mean of disparities. Finally, the Rao-Stirling index combines into one score the variety, disparity, and also balance, as a third aspect of diversity. The Rao-Stirling index was calculated by the default method implemented in the *diverse* package [56]. All measures were transformed according to Yegros-Yegros et al. [45] based on the empirically observed minimum and maximum within a range from 0 to 1 (i.e., $x' = (x - \min(x))/(\max(x) - \min(x))$).

2.2.3. Idea Density

In previous empirical work and theory, idea density has also been tied to creativity [48]. Propositional idea density can be approximated by the number of verbs, adjectives, adverbs, prepositions, and conjunctions divided by the total number of words [59,60]. In this work, we used an automatic scoring of idea density, which refines the proposition count (and sometimes the word count) by means

of adjustment rules. For example, noun phrases introduced by a copula are counted as proposition, but adjective phrases introduced by a copula are not (for more rules see Brown et al. [59]). For each of the datasets, we calculated two indicators of idea density: first, we used the idea density of a publication's title, and second, we used the idea density of a publication's abstract. Idea density was calculated by means of the software CPIDR 3.2 (<http://ai1.ai.uga.edu/caspr/>; [59]). CPIDR 3.2 displayed strong validity findings, as automatic scores correlated close to unity with human raters ($r = 0.97$). Initial validity findings of idea density as an indicator of creative units of meaning were provided by Runco et al. [48].

2.2.4. Originality

Shibayama and Wang [39] found the highest correlation between originality and self-assessed theoretical originality when the originality measure was weighted according to the number of references in citing works ($\text{Orig}_{\text{weighted1}}$; r was slightly below 0.20 in that case). However, when applying their originality measures to documents outside the life sciences, they suggest checking the optimal number of papers belonging to the citing set. Hence, we decided to calculate $\text{Orig}_{\text{base}}$ and $\text{Orig}_{\text{weighted1}}$ when four documents ($S = 4$) made up the citing set and when eight documents ($S = 8$) made up the citing set ($\text{Orig}_{\text{weighted2}}$ was not calculated for this study). Citation counts for all references in the target articles were not easily available for all of the works, so we decided to omit the weighted2 formula from Shibayama and Wang [39]. Further, Shibayama-Wang originality was only calculated when at least 50% of the citing articles and when at least S citing articles were available in the retrieved data.

2.3. Data Analysis

The full analysis was conducted with the statistical software R [61] and its bibliometrix package [49]. Several of the indicators used had strong methodological overlap (i.e., identical terms in the respective formulas). In addition, some missing data patterns were Missing at Random (MAR; e.g., missing values on $\text{Orig}_{\text{base}}$ with eight articles in the citing set depended on publication year, because some of the more recent articles had not yet received eight citations), so we performed estimations by means of multiple imputation in the R package mice [62]. As suggested in the literature, we used $m = 40$ imputed datasets [63]. Thus, we decided to run parts of the analyses separately. That is, for the validity analyses of idea density and the openness indicators, all variants of Shibayama-Wang originality were analyzed separately to prevent correlation estimates that were inconsistent with the fact that these variables had strong methodological overlap (resulting from the respective formulas). Correlations were estimated by means of the R package brms [64] in order to use Bayesian statistical inference as a solution to known problems inherent in the practice of null hypothesis testing [65]. In particular, we report 95% credible intervals and examined the posterior probabilities for the hypothesis that the correlations were greater than zero by means of the hypothesis() function in brms. Shibayama-Wang originality was used as a criterion for convergent validity, whereas impact was used as a criterion for discriminant validity. Finally, exploratory graph analysis (EGA) [66,67], as implemented in the R package EGAnet [68], was used to explore the clustering of the openness and idea density indicators as part of the convergent validity analyses. EGA was applied to the correlation matrix obtained from the above-described multiple imputation procedure and the harmonic mean of the available datapoints to accurately reflect the amount of missing values in the data (see Table 1 and below for the found pattern of missing values).

Analysis scripts, data to reproduce the analyses presented in this work, and the fitted brms model objects are available in an Open Science Framework repository: <https://osf.io/2qnvv/>.

3. Results

3.1. Missing Data Pattern and Descriptive Statistics

Descriptive statistics for both the creativity and bibliographic datasets are depicted in Table 1. First, it is useful to look at the missing data patterns. When controlling for publication year, a few datapoints (creativity dataset: 1.3%; bibliometrics dataset: 0.8%) were lost for analyses because

of missing values for publication year in both datasets. The openness and idea density indicators displayed small to moderate amounts of missing data (creativity dataset: maximally 6.1%; bibliometrics dataset: maximally 19.93%). Notably, the idea density of the titles was fully available for all articles in both datasets. As expected, the amount of missing data increased substantially for Shibayama-Wang originality. To further illustrate the assumed mechanism via publication year, we predicted the amount of missing data in a Bayesian logistic regression by publication year. For both datasets, the amount of missing data with $S = 4$ was significantly predicted by publication year (creativity dataset: odds ratio = 1.55, 95% credible interval: [1.48, 1.63]; bibliometrics dataset: odds ratio = 1.29, 95% credible interval: [1.26, 1.32]). The expected MAR pattern also fit with the observation that the amount of missing data increased further when using Shibayama-Wang originality based on $S = 8$ as compared to $S = 4$, because all missing values in the $S = 4$ scores were also missing in the $S = 8$ scores in both datasets (i.e., a strictly monotone pattern of missing values). However, missing values of all other study variables were in accordance with a Missing Completely at Random (MCAR) pattern for the creativity dataset (Hawkin's Test: $p < 0.001$, non-parametric test of homoscedasticity: $p = 0.231$), but not for the bibliometric dataset (Hawkin's Test: $p < 0.001$, non-parametric test of homoscedasticity: $p < 0.001$). Given that we are not aware of any comparable Bayesian approach, this was tested by means of Jamshidian and Jalal's [69] two-step procedure, as implemented in the R package MissMech [70]. The non-random pattern of missing values for the bibliometric dataset emerged from only the openness indicators and, hence, most likely a Missing at Random mechanism was responsible for missing data here. Hence, our analytical strategy to use multiple imputation for statistical inference and estimation of correlations between all study variables and Shibayama-Wang originality separately for all Shibayama-Wang originality variants was largely justified.

Both datasets were largely comparable with respect to total citations, though the standard deviation for the bibliometrics dataset was larger than the creativity dataset. All openness indicators were found to be larger on average (and had slightly larger standard deviations) for the creativity dataset, whereas the idea density indicators were highly comparable between datasets on average and with respect to standard deviations (see Table 1). Average Shibayama-Wang originalities were all smaller for the creativity dataset as compared to the bibliometrics dataset, but the standard deviations of Shibayama-Wang originalities were all higher for the creativity dataset than in the bibliometrics dataset.

Table 1. Descriptive statistics.

	Creativity Dataset			Bibliometric Dataset		
	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>
Total citations	1643	11.31	17.22	2986	11.05	30.55
Total citations controlled	1622	0.00	14.94	2962	0.00	29.09
Variety—Journals	1633	0.26	0.13	2941	0.08	0.07
Disparity—Journals	1583	0.74	0.11	2624	0.87	0.10
Rao-Stirling—Journals	1583	0.06	0.07	2624	0.01	0.03
Variety—First authors	1630	0.21	0.11	2911	0.06	0.05
Disparity—First authors	1543	0.95	0.09	2391	0.96	0.11
Rao-Stirling—First authors	1543	0.03	0.07	2391	0.01	0.04
Idea density—Abstracts	1631	0.30	0.04	2942	0.30	0.04
Idea density—Titles	1643	0.27	0.11	2986	0.24	0.09
Orig _{base} - $S = 4$	953	0.93	0.15	1360	0.94	0.14
Orig _{weighte1} - $S = 4$	953	0.96	0.10	1360	1.00	0.03
Orig _{base} - $S = 8$	616	0.93	0.12	912	0.95	0.12
Orig _{weighte1} - $S = 8$	616	0.95	0.11	912	0.99	0.06

S = set size of citing articles on which Shibayama-Wang originality was based.

3.2. Validity Findings

3.2.1. Discriminant Validity Findings

To achieve discriminant validity, measures of independent attributes should be uncorrelated with one another, or at least less strongly correlated than when compared to convergent validity correlations (see below) with the target measure (here, the various indicators of creative quality). Impact, as measured by total citations, was used as a criterion for discriminant validity. The correlations between citations (corrected for publication year) and various creative quality indicators are depicted in Table 2 for the creativity dataset. All openness indicators—with disparities as exceptions—displayed small positive (r s ranged from 0.103 to 0.157) correlations with the number of citations. For all variety and Rao-Stirling measures the 95% credible intervals did not cover a value of zero and the posterior probabilities for the hypothesis that the correlation is larger than zero was found to be larger than 0.95. Both idea density indicators displayed negligible correlations (95% credible intervals covered zero for both correlations; see Table 2) with the number of citations for the creativity dataset. These findings were replicated for the bibliometrics dataset (see Table 3) with positive correlations between citations and only variety measures as openness indicators (r s ranged from 0.090 to 0.097). In addition, negligibly small correlations between citations and all other indicators were found, as was indicated by 95% credible intervals and also examination of posterior probabilities of the one-sided hypothesis that the correlation is greater zero. Also, Shibayama-Wang originality had only negligible correlations with impact across datasets (see Tables 2 and 3), which adds to the validity evidence of these measures.

3.2.2. Convergent Validity Findings

The various indicators of creative quality were expected to be positively and moderately inter-correlated as compared to the correlations between these indicators and impact. This analysis can also be understood as a test of concurrent *criterion validity*. The main criteria for convergent validity were the variants of the recently validated Shibayama-Wang originality measures [39]. The convergent validity results for the creativity dataset revealed that the exact formula used for Shibayama-Wang originality ($Orig_{base}$ vs. $Orig_{weighted1}$) was more influential on convergent validity correlations than was the size of the citing set S (four vs. eight). Correlations between varieties and Shibayama-Wang originality were all positive (r s ranged from 0.033 to 0.221) and in almost all of these cases credible intervals did not cover zero. In particular, when the correction for the number of references in the citing articles was applied, correlations were clearly higher as compared to the correlations without this correction (see Table 2). In addition, correlations between all other openness indicators and Shibayama-Wang originality were negligible (see Table 2), with only the Rao-Stirling based on journals as an exception (see Table 2). Convergent validity findings of the idea density indicators for the creativity dataset further revealed that a small positive and significant correlation was found only for idea density of abstracts for almost all variants of Shibayama-Wang originality (see Table 2). Finally, convergent validity of the various indicators for openness and idea density were analyzed by means of exploratory graph analysis to reveal their underlying dimensionality. Three clusters emerged for these indicators in the creativity dataset (see Figure 1). The first cluster (red nodes in Figure 1) comprised the disparity based on journals and Rao-Stirling openness measures, the second cluster comprised both variety openness measures, and the third cluster comprised both idea density indicators (please note that disparity based on first authors did not cluster with any of the other variables).

Table 2. Validity findings for the creativity dataset.

	Validity Criterion Total Citations Controlled	Orig _{base} - S = 4	Orig _{weighted1} - S = 4	Orig _{base} - S = 8	Orig _{weighted1} - S = 8
Total citations controlled	-	-0.023 [-0.075, 0.029]	0.036 [-0.018, 0.090]	-0.044 [-0.103, 0.017]	0.002 [-0.059, 0.064]
Variety—Journals	0.157 [0.109, 0.203] *	0.062 [-0.001, 0.122] *	0.186 [0.131, 0.236] *	0.033 [-0.041, 0.112]	0.189 [0.118, 0.262] *
Disparity—Journals	-0.027 [-0.076, 0.023]	-0.039 [-0.104, 0.024]	-0.066 [-0.132, -0.003]	0.035 [-0.046, 0.109]	0.025 [-0.070, 0.131]
Rao-Stirling—Journals	0.135 [0.086, 0.184] *	0.002 [-0.062, 0.062]	0.075 [0.017, 0.129] *	-0.035 [-0.124, 0.044]	0.057 [-0.018, 0.123]
Variety—First authors	0.154 [0.107, 0.201] *	0.077 [0.013, 0.140] *	0.204 [0.153, 0.253] *	0.064 [-0.006, 0.131] *	0.221 [0.153, 0.282] *
Disparity—First authors	-0.014 [-0.063, 0.035]	-0.005 [-0.065, 0.056]	0.013 [-0.046, 0.081]	-0.003 [-0.067, 0.063]	0.022 [-0.060, 0.149]
Rao-Stirling—First authors	0.103 [0.051, 0.154] *	-0.031 [-0.092, 0.028]	-0.001 [-0.096, 0.061]	-0.077 [-0.150, -0.005]	-0.062 [-0.210, 0.033]
Idea density—Abstracts	0.025 [-0.023, 0.073]	0.062 [-0.004, 0.124] *	0.058 [-0.001, 0.115] *	0.088 [0.017, 0.162] *	0.062 [-0.040, 0.145]
Idea density—Titles	-0.006 [-0.054, 0.042]	-0.016 [-0.084, 0.044]	-0.014 [-0.074, 0.052]	-0.014 [-0.083, 0.055]	-0.051 [-0.133, 0.029]

S = set size of citing articles on which Shibayama-Wang originality was based. 95% credible intervals are provided along with the correlation estimates, which were combined across $m = 40$ imputed datasets by means of the brms function `brm_multiple()`. Multiple imputation was carried out by means of the R package `mice` [62]. * posterior probability for the hypothesis that the correlation is greater than zero exceeded 0.95.

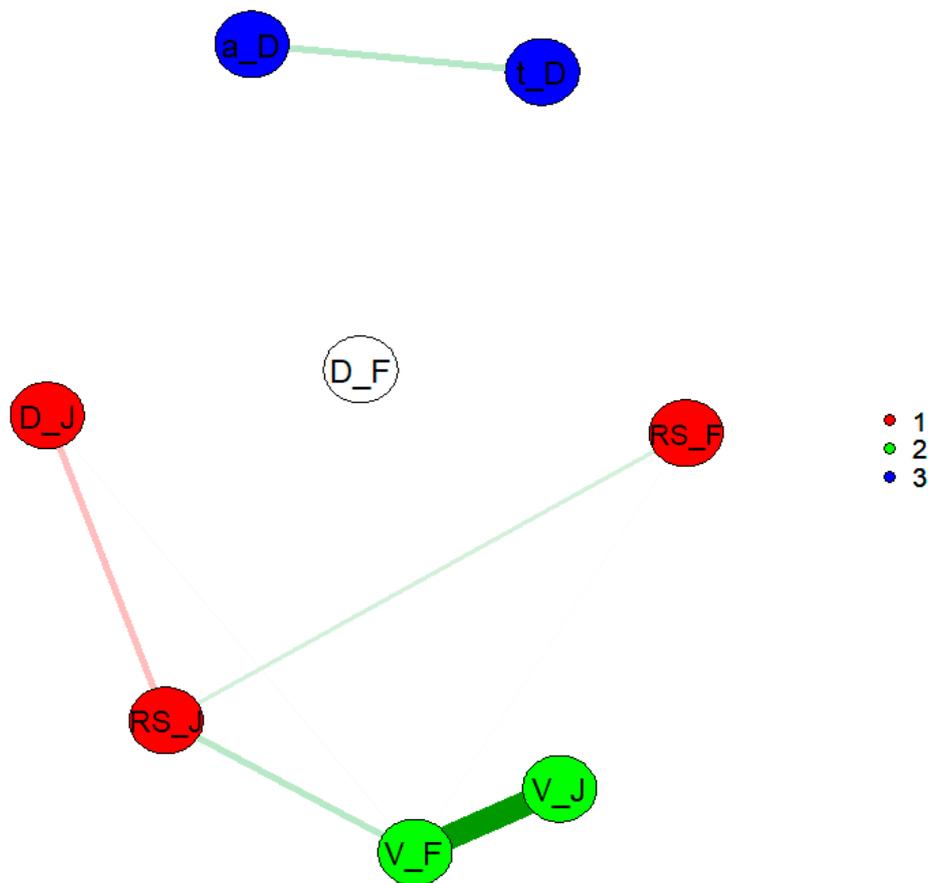


Figure 1. Exploratory graph analysis of the openness and idea density indicators for the creativity dataset. V_J = variety based on journals; D_J = disparity based on journals; RS_J = Rao-Stirling based on journals; V_F = variety based on first authors; D_F = disparity based on first authors; RS_F = Rao-Stirling based on first authors; a_D = idea density based on abstracts; t_D = idea density based on titles.

Unlike the discriminant validity findings, the convergent validity findings obtained for the creativity dataset did not comparably replicate for the bibliometrics dataset (see Table 3). There were mostly negligible correlations between the creative quality indicators and Shibayama-Wang originality (see Table 3). The only exception was variety based on referenced journals. This openness indicator displayed small positive and significant correlations with all variants of Shibayama-Wang originality. This relationship was slightly stronger when the correction for the number of references in the citing articles was applied (and highest when S was 8; see Table 3). In addition, correlations between variety

based on first authors and Shibayama-Wang originality had only credible intervals not covering zero and remarkably large posterior probabilities for the assumed hypothesis, when Shibayama-Wang originality was calculated without correction (see Table 3). Furthermore, the clustering of openness and idea density was somewhat different when compared to the creativity dataset (see Figure 2). Again, the idea density indicators clustered together (blue nodes), but their inter-relatedness ($r = 0.108$, 95% credible interval: [0.072, 0.144]) was lower than in the creativity dataset ($r = 0.264$, 95% credible interval: [0.220, 0.307]). Nonetheless, the strong relationship between both varieties replicated well; yet, disparity and Rao-Stirling measures based on journals in the bibliometric dataset did not cluster together with the variety measures. Hence, the bibliometrics dataset resulted in only two clusters, whereas the creativity dataset had three.

Table 3. Validity findings for the bibliometric dataset.

Validity Criterion	Total Citations Controlled	Orig _{base} - S = 4	Orig _{weighte1} - S = 4	Orig _{base} - S = 8	Orig _{weighte1} - S = 8
Total citations controlled	–	–0.047 [–0.086, –0.007]	–0.001 [–0.044, 0.044]	–0.039 [–0.082, 0.004]	0.006 [–0.038, 0.051]
Variety—Journals	0.097 [0.061, 0.132] *	0.066 [0.009, 0.111] *	0.081 [0.031, 0.128] *	0.099 [0.046, 0.150] *	0.121 [0.059, 0.173] *
Disparity—Journals	0.014 [–0.023, 0.050]	–0.087 [–0.144, –0.025]	–0.043 [–0.123, 0.131]	–0.099 [–0.163, –0.040]	–0.103 [–0.182, –0.022]
Rao-Stirling—Journals	0.016 [–0.020, 0.051]	0.015 [–0.033, 0.055]	–0.013 [–0.143, 0.043]	–0.017 [–0.034, 0.060]	–0.083 [–0.333, 0.044]
Variety—First authors	0.090 [0.055, 0.126] *	0.069 [0.022, 0.110] *	0.043 [–0.107, 0.099]	0.091 [0.044, 0.137] *	0.036 [–0.112, 0.129]
Disparity—First authors	0.025 [–0.016, 0.063]	–0.042 [–0.083, 0.003]	0.006 [–0.051, 0.151]	–0.024 [–0.073, 0.048]	0.013 [–0.061, 0.246]
Rao-Stirling—First authors	0.026 [–0.014, 0.087]	–0.007 [–0.052, 0.037]	–0.027 [–0.303, 0.043]	–0.037 [–0.092, 0.018]	–0.022 [–0.182, 0.041]
Idea density—Abstracts	–0.011 [–0.046, 0.025]	0.012 [–0.048, 0.070]	–0.006 [–0.074, 0.070]	0.028 [–0.052, 0.136]	–0.063 [–0.158, 0.064]
Idea density—Titles	0.004 [–0.032, 0.039]	0.037 [–0.019, 0.091]	–0.022 [–0.085, 0.047]	0.013 [–0.063, 0.092]	0.004 [–0.059, 0.083]

S = set size of citing articles on which Shibayama-Wang originality was based. 95% credible intervals are provided along with the correlation estimates which are combined across $m = 40$ imputed datasets by means of the brms function `brm_multiple()`. Multiple imputation was carried out by means of the R package `mice` [62]. * posterior probability for the hypothesis that the correlation is greater than zero exceeds 0.95.

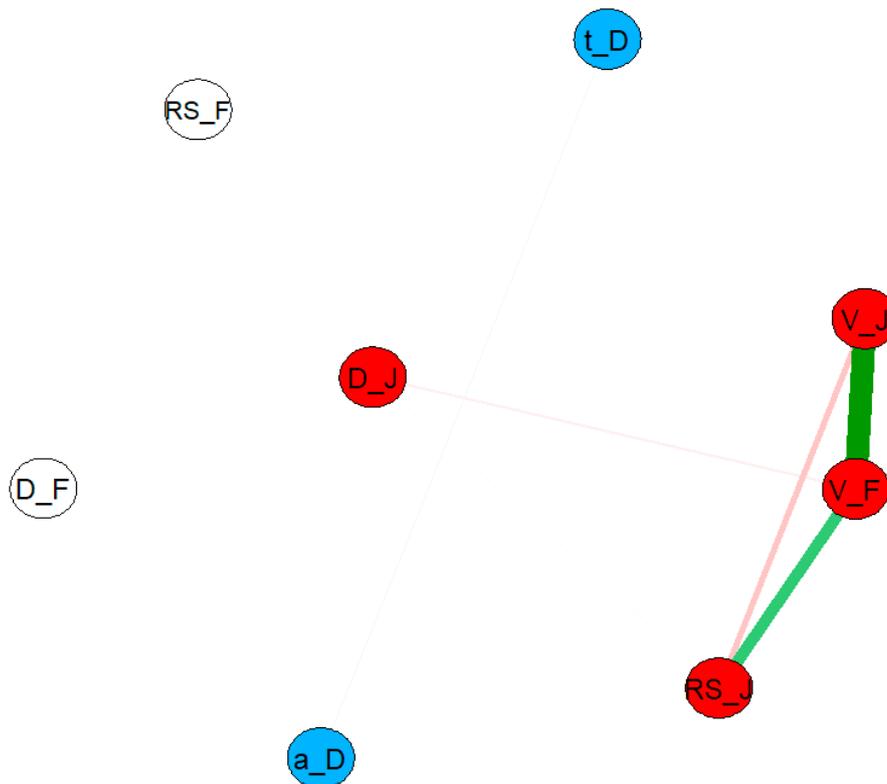


Figure 2. Exploratory graph analysis of the openness and idea density indicators for the bibliometric dataset. V_J = variety based on journals; D_J = disparity based on journals; RS_J = Rao-Stirling based on journals; V_F = variety based on first authors; D_F = disparity based on first authors; RS_F = Rao-Stirling based on first authors; a_D = idea density based on abstracts; t_D = idea density based on titles.

4. Discussion

In general, objectively quantifying the creative qualities of scientific work (e.g., originality) is not an easy undertaking [12,39], but we predicted that openness and idea density, both theoretically connected with the creativity of scientific work, would be informative. With this in mind, we examined a number of indicators, including openness, based on two sources of information (i.e., the journals in which referenced works were published and the authors of the referenced work), three commonly applied diversity bibliographic indices (i.e., variety, disparity, and the Rao-Stirling index), and idea density, based on abstracts and titles of a scientific work. These indicators were correlated both with impact and with Shibayama and Wang's [39] recently validated measure of originality. Impact was operationally defined as the total number of citations received.

Some researchers have used impact as a proxy for creativity [25–27,29,30], but impact and originality can be distinguished on theoretical grounds [14–18]. Hence, we used citations to test discriminant validity, the expectation being that originality and the other creative scholarship indicators should be unrelated to impact, or perhaps only very slightly related to it. We employed two datasets to ensure generalizability of the findings. In addition, Shibayama-Wang [39] originality measures were used as criteria for convergent validity of the various proposed openness and idea density indicators. Finally, the nomological net of these creative quality indicators was explored.

Idea density indicators displayed better discriminant validity as compared to the openness indicators. This was indicated by negligibly sized correlations between all idea density indicators and impact across both studied datasets. For openness indicators, however, small positive correlations were consistently found. The sizes of the correlations between openness and impact imply that maximally $\approx 2.5\%$ of the variation is shared by these variables. Thus, despite the overlap, these findings are still in accordance with discriminant validity (i.e., unique variation in both openness and impact was much larger than their shared variation).

Results across the datasets were less consistent in terms of convergent validity. For the creativity dataset, variety correlated positively and more strongly with Shibayama-Wang originality when it was corrected for the number of references included in the works of the citing set. However, the largest correlation implied a shared variation of approximately 4.9%. Hence, given that openness correlated mostly equally strongly with the criterion for discriminant validity and the criterion for convergent validity, one must conclude that there is no clear evidence of convergent validity for openness [71]. Yet, it should be noted that correlations between originality and scientists' subjective ratings were found to be comparable in size by Shibayama and Wang [39]; hence, the validity results of variety as an openness indicator can also be considered as good as previous validity findings reported for originality of scientific work. Finally, idea density was found to correlate significantly and positively with Shibayama-Wang originality under very specific conditions, implying that this aspect of creative quality seems to be rather different from impact, openness, and Shibayama-Wang originality.

The latter observation regarding idea density was also found for the bibliometrics dataset. However, for this dataset the varieties did not correlate as consistently with Shibayama-Wang originality as they did in the creativity dataset. Only the variety of journals correlated positively with all originality measures. Hence, in this study, the variety of cited journals was found to be consistently positively related with both impact and Shibayama-Wang originality across datasets.

The clustering of all openness and idea density indicators was further checked (see Figures 1 and 2). Here, the idea density indicators built one cluster (with a stronger relationship between abstracts and titles for the creativity dataset as compared to the bibliometrics dataset). Moreover, both variety measures for openness clustered consistently together across datasets. Given that both measures were based on the same references, this could have been expected. Then again, the Rao-Stirling indices were also based on the same references, but they did not cluster consistently across datasets: for the creativity dataset both Rao-Stirling indices clustered together separately from variety, whereas for the bibliometrics dataset their separate cluster collapsed.

This research was predicated on theories from the field of creativity research, which holds that impact is distinct from originality, and both are distinct from creativity [14–18]. While there might be overlap, the overlap between impact and originality would be minimal. This follows from the fact that originality is apparent when an idea (expressed in a publication, concept, or method) differs from what came before it. In the sciences, though, this is probably more a matter of degree than an all-or-nothing result, which is why innovation can be described as either incremental or radical [72], where incremental innovation represents a small step forward, (i.e., a small difference between what existed earlier and the new innovation), and radical innovation represents a large step forward (i.e., a large discrepancy between what existed earlier and the new innovation). Admittedly, all innovations are new, and we are aware of the irony of using a dichotomy (incremental vs. radical) in an argument that suggests creativity should be viewed as continuous rather than dichotomous. The key point is that originality must be determined by comparing a new idea with what previously existed. Impact, on the other hand, is determined by comparing a new idea with what follows. If a new idea has a strong impact, what follows (e.g., citations) will be influenced. In short, originality is assessed by looking backward, while impact is assessed by looking forward.

Impact and originality are distinct, as described above, but both may be related to creativity. For example, originality is a prerequisite for creativity, but originality alone does not fully encompass creativity; creativity also requires effectiveness [20,21]. In the sciences, effectiveness might be implied by impact [30], meaning that other indicators could potentially be constructed. Hence, bibliometric data might harbor indicators for both creativity requirements and should be examined in the future.

Limitations and Future Directions

The study is limited to the research fields studied as well as to the creative scholarship indicators used, which include idea density (of titles and abstracts), openness (i.e., diversity of journals cited and diversity of authors cited), and originality. Not all variants of Shibayama-Wang originality were used, for reasons described in the Methods. Finally, this study depended on the WoS database, which is extensive but still may be considered a limitation.

It should further be noted that for the evaluation of early career scientists the studied indicators in this work look promising. However, the analyses presented here are limited to articles as level of analysis and, hence, it remains an open empirical question how individual differences between scientists in these indicators might be helpful in grant decisions, for example. In this vein, the relatedness of openness and idea density indicators with independence indicators, as suggested by van den Besselaar et al. [12], might provide a multidimensional set of quality indicators to complement the information provided by impact indicators. Clearly, more research is needed in this regard.

Some of the indicators in this study (variety, disparity, and Rao-Stirling indices) were transformed into the range from zero to one simply for the reason to have all studied indicators in the same range. These transformed variables did not affect the correlational findings presented here, but it has been pointed out that diversity measures for bibliometric studies should be interpretable as percentages or ratios [58]. Transformations can prevent such an interpretation and, hence, their use should be carefully considered in future work.

The method introduced here can be extended in future research. One such investigation might use the idea from Shibayama and Wang [39] where originality is determined by contrasting the references of a target paper with the references of papers that cite the target paper; this type of analysis could be done with keywords, or perhaps even titles. It would also be promising to examine the proposed indicators of creative quality at the author level to assess their relationship with other author-level quantifications of quality, such as the independence index [12]. A final promising line of research would be to use all reliable indicators together, in an optimized equation, for a multivariate and, thus, a more comprehensive bibliometric explanation of creative scholarship.

5. Conclusions

This work extends previous lines of research, attempting to quantify the creative quality of scholarly work. Importantly, empirical findings on the inter-relatedness of creative scholarship indicators of scientific work are still scarce [12,39], and the present work addresses this gap in the literature. Several theoretical arguments to distinguish between impact and creative qualities have been proposed in the literature and in the current work, and the current findings related to discriminant validity support this view. In terms of convergent validity, however, correlations across the different indicators (openness, idea density, and originality) did not notably exceed the highest correlations between impact and some of those indicators (i.e., variety of cited first authors), but the highest found correlations were of comparable size ($r \approx 0.20$) to those reported by Shibayama and Wang [39]. Notably, correlations in the range $0.10 \leq r < 0.30$ are considered small effect sizes in the social sciences literature [73]. Importantly, some caution when studying these measures in future research is needed, because various measures were based on the same source of information (e.g., the cited references) or shared terms used in the formula to calculate them (e.g., Rao-Stirling combines variety and disparity with balance in a single score). However, idea density based on titles and abstracts are much less affected by such technical issues, rendering these measures promising despite their found unrelatedness to other creative quality measures. Importantly, this work, as a thorough investigation of the nomological net of several creative quality indicators, provides a good starting point for related future research.

Author Contributions: Conceptualization, M.A.R.; methodology, M.A.R. and B.F.; software, B.F.; validation, B.F.; investigation, B.F.; data curation, B.F.; writing—original draft preparation, M.A.R. and B.F.; writing—review and editing, M.A.R. and B.F.; visualization, B.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bornmann, L. Scientific peer review. *Annu. Rev. Inf. Sci.* **2011**, *45*, 197–245. [[CrossRef](#)]
2. Lawani, S.M. Some bibliometric correlates of quality in scientific research. *Scientometrics* **1986**, *9*, 13–25. [[CrossRef](#)]
3. Freese, J.; Peterson, D. Replication in social science. *Annu. Rev. Sociol.* **2017**, *43*, 147–165. [[CrossRef](#)]
4. McNutt, M. Journals unite for reproducibility. *Science* **2014**, *346*, 679. [[CrossRef](#)] [[PubMed](#)]
5. Hartley, J. Authors and their citations: A point of view. *Scientometrics* **2017**, *110*, 1081–1084. [[CrossRef](#)]
6. Garfield, E. Citation indexes for science. *Science* **1955**, *122*, 108–111. [[CrossRef](#)] [[PubMed](#)]
7. Garfield, E. Citation analysis as a tool in journal evaluation. *Science* **1972**, *178*, 471–479. [[CrossRef](#)]
8. Pan, R.K.; Fortunato, S. Author Impact Factor: Tracking the dynamics of individual scientific impact. *Sci. Rep.* **2015**, *4*, 4880. [[CrossRef](#)] [[PubMed](#)]
9. Heinze, T.; Bauer, G. Characterizing creative scientists in nano-S&T: Productivity, multidisciplinary, and network brokerage in a longitudinal perspective. *Scientometrics* **2007**, *70*, 811–830. [[CrossRef](#)]
10. Glänzel, W.; Moed, H. Journal impact measures in bibliometric research. *Scientometrics* **2002**, *53*, 171–193. [[CrossRef](#)]
11. Shu, F.; Quan, W.; Chen, B.; Qiu, J.; Sugimoto, C.R.; Larivière, V. The role of Web of Science publications in China's tenure system. *Scientometrics* **2020**, 1–13. [[CrossRef](#)]
12. Van den Besselaar, P.; Sandström, U.; van der Weijden, I. The independence indicator. In *Science & Technology Indicators*; Archambault, E., Gingras, Y., Larivière, V., Eds.; OST & Science Metrix: Montreal, QC, Canada, 2012; pp. 131–141.
13. Brown, L.D.; Gardner, J.C. Using citation analysis to assess the impact of journals and articles on contemporary accounting research (CAR). *J. Acc. R.* **1985**, *23*, 84–109. [[CrossRef](#)]
14. Albert, R.S. Madison Avenue comes to academe. *Creat. Res. J.* **1995**, *8*, 427–429. [[CrossRef](#)]
15. Gardner, H.; Policastro, E. Naive Judgment and Expert Assessment: A Critique of the Attributional Perspective. *Creat. Res. J.* **1995**, *8*, 391–395. [[CrossRef](#)]

16. Ludwig, A.M. What “explaining creativity” doesn’t explain. *Creat. Res. J.* **1995**, *8*, 413–416. [[CrossRef](#)]
17. Runco, M.A. Insight for creativity, expression for impact. *Creat. Res. J.* **1995**, *8*, 377–390. [[CrossRef](#)]
18. Sternberg, R.J. If you change your name to Mark Twain, will you be judged as creative? *Creat. Res. J.* **1995**, *8*, 367–370. [[CrossRef](#)]
19. Moed, H.F. *Applied Evaluative Informetrics*; Springer: Cham, Switzerland, 2017.
20. Runco, M.A.; Jaeger, G.J. The standard definition of creativity. *Creat. Res. J.* **2012**, *24*, 92–96. [[CrossRef](#)]
21. Heinze, T. Creative accomplishments in science: Definition, theoretical considerations, examples from science history, and bibliometric findings. *Scientometrics* **2013**, *95*, 927–940. [[CrossRef](#)]
22. Lubart, T.; Guignard, J. The generality-specificity of creativity: A multivariate approach. In *Creativity: From Potential to Realization*; Sternberg, R.J., Grigorenko, E.L., Singer, J.L., Eds.; American Psychological Association: Washington, DC, USA, 2004; pp. 43–56. [[CrossRef](#)]
23. Sternberg, R.J.; Lubart, T. *Defying the Crowd: Cultivating Creativity in an Age of Conformity*; Free Press: New York, NY, USA, 1995.
24. Runco, M.A.; Pritzker, S. *Encyclopedia of Creativity*, 3rd ed.; Elsevier: San Diego, CA, USA, In press.
25. Simonton, D.K. *Scientific Genius: A Psychology of Science*; Cambridge University Press: New York, NY, USA, 1988.
26. Simonton, D.K. *Creativity in Science: Chance, Logic, Genius, and Zeitgeist*; Cambridge University Press: New York, NY, USA, 2004. [[CrossRef](#)]
27. Forthmann, B.; Szardenings, C.; Dumas, D. Testing equal odds in creativity research. *Psychol. Aesthet. Creat. Arts.* in press. [[CrossRef](#)]
28. Forthmann, B.; Szardenings, C.; Dumas, D. On the conceptual overlap between the fluency contamination effect in divergent thinking scores and the chance view on scientific creativity. *J. Creat. Behav.* in press. [[CrossRef](#)]
29. Forthmann, B.; Leveling, M.; Dong, Y.; Dumas, D. Investigating the quantity–quality relationship in scientific creativity: an empirical examination of expected residual variance and the tilted funnel hypothesis. *Scientometrics*. in press. [[CrossRef](#)]
30. Wang, J. Knowledge creation in collaboration networks: Effects of tie configuration. *Res. Policy* **2016**, *45*, 68–80. [[CrossRef](#)]
31. Franceschet, M.; Costantini, A. The effect of scholar collaboration on impact and quality of academic papers. *J. Inf.* **2010**, *4*, 540–553. [[CrossRef](#)]
32. Leimu, R.; Koricheva, J. Does scientific collaboration increase the impact of ecological articles? *BioScience* **2005**, *55*, 438–443. [[CrossRef](#)]
33. Long, H. More than appropriateness and novelty: Judges’ criteria of assessing creative products in science tasks. *Think. Ski. Creat.* **2014**, *13*, 183–194. [[CrossRef](#)]
34. Azoulay, P.; Zivin, J.S.G.; Manso, G. Incentives and creativity: Evidence from the academic life sciences. *Rand J. Econ.* **2011**, *42*, 527–554. [[CrossRef](#)]
35. Foster, J.G.; Rzhetsky, A.; Evans, J.A. Tradition and innovation in scientists’ research strategies. *Am. Sociol. Rev.* **2015**, *80*, 875–908. [[CrossRef](#)]
36. Arora, S.K.; Porter, A.L.; Youtie, J.; Shapira, P. Capturing new developments in an emerging technology: An updated search strategy for identifying nanotechnology research outputs. *Scientometrics* **2013**, *95*, 351–370. [[CrossRef](#)]
37. Funk, R.J.; Owen-Smith, J. A dynamic network measure of technological change. *Manag. Sci.* **2017**, *63*, 791–817. [[CrossRef](#)]
38. Wu, L.; Wang, D.; Evans, J.A. Large teams develop and small teams disrupt science and technology. *Nature* **2019**, *566*, 378–382. [[CrossRef](#)] [[PubMed](#)]
39. Shibayama, S.; Wang, J. Measuring originality in science. *Scientometrics* **2020**, *122*, 409–427. [[CrossRef](#)]
40. Anastasi, A. *Psychological Testing*, 6th ed.; Prentice Hall International: New York, NY, USA, 1988.
41. Dollinger, S.J.; Urban, K.K.; James, T.A. Creativity and openness: Further validation of two creative product measures. *Creat. Res. J.* **2004**, *16*, 35–47. [[CrossRef](#)]
42. Feist, G.J. A meta-analysis of personality in scientific and artistic creativity. *Pers. Soc. Psychol. Rev.* **1998**, *2*, 290–309. [[CrossRef](#)] [[PubMed](#)]
43. McCann, S.J. Conservatism, openness, and creativity: Patents granted to residents of American states. *Creat. Res. J.* **2011**, *23*, 339–345. [[CrossRef](#)]

44. Stirling, A. A general framework for analysing diversity in science, technology and society. *J. R. Soc. Interface* **2007**, *4*, 707–719. [[CrossRef](#)]
45. Yegros-Yegros, A.; Rafols, I.; D’Este, P. Does interdisciplinary research lead to higher citation impact? The different effect of proximal and distal interdisciplinarity. *PLoS ONE* **2015**, *10*, e0135095. [[CrossRef](#)]
46. Runco, M.A.; Acar, S. Divergent Thinking. In *The Cambridge Handbook of Creativity*; Kaufman, J.C., Sternberg, R.J., Eds.; Cambridge University Press: New York, NY, USA, 2019; pp. 224–253.
47. Forthmann, B.; Jendryczko, D.; Scharfen, J.; Kleinkorres, R.; Benedek, M.; Holling, H. Creative ideation, broad retrieval ability, and processing speed: A confirmatory study of nested cognitive abilities. *Intelligence* **2019**, *75*, 59–72. [[CrossRef](#)]
48. Runco, M.A.; Turkman, B.; Acar, S.; Nural, M.V. Idea density and the creativity of written works. *J. Genius Eminence* **2017**, *2*, 26–31. [[CrossRef](#)]
49. Aria, M.; Cuccurullo, C. bibliometrix: An R-tool for comprehensive science mapping analysis. *J. Inf.* **2017**, *11*, 959–975. [[CrossRef](#)]
50. Melo, A.S.; Bini, L.M.; Carvalho, P. Brazilian articles in international journals on Limnology. *Scientometrics* **2006**, *67*, 187–199. [[CrossRef](#)]
51. Rubenson, D.L.; Runco, M.A. The psychoeconomic approach to creativity. *New Ideas Psychol.* **1992**, *10*, 131–147. [[CrossRef](#)]
52. Root-Bernstein, R.S. Setting the stage for discovery. *Science* **1988**, *28*, 26–34. [[CrossRef](#)]
53. Rubenson, D.L.; Runco, M.A. The psychoeconomic view of creative work in groups and organizations. *Creat. Innov. Manag.* **1995**, *4*, 232–241. [[CrossRef](#)]
54. Mednick, S. The associative basis of the creative process. *Psychol. Rev.* **1962**, *69*, 220–232. [[CrossRef](#)] [[PubMed](#)]
55. Beketayev, K.; Runco, M.A. Scoring divergent thinking tests with a semantics-based algorithm. *Eur. J. Psychol.* **2016**, *12*, 210–220. [[CrossRef](#)]
56. Guevara, M.R.; Hartmann, D.; Mendoza, M. Diverse: An R Package to analyze diversity in complex systems. *R J.* **2016**, *8*, 60–78. [[CrossRef](#)]
57. Schubert, A. A reference-based Hirschian similarity measure for journals. *Scientometrics* **2010**, *84*, 133–147. [[CrossRef](#)]
58. Zhang, L.; Rousseau, R.; Glänzel, W. Diversity of references as an indicator of the interdisciplinarity of journals: Taking similarity between subject fields into account. *J. Assoc. Inf. Sci. Technol.* **2016**, *67*, 1257–1265. [[CrossRef](#)]
59. Brown, C.; Snodgrass, T.; Kemper, S.J.; Herman, R.; Covington, M.A. Automatic measurement of propositional idea density from part-of-speech tagging. *Behav. Res. Methods* **2008**, *40*, 540–545. [[CrossRef](#)]
60. Snowdon, D.A.; Kemper, S.J.; Mortimer, J.A.; Greiner, L.H.; Wekstein, D.R.; Markesbery, W.R. Linguistic ability in early life and cognitive function and Alzheimer’s disease in late life: Findings from the Nun Study. *JAMA* **1996**, *275*, 528–532. [[CrossRef](#)]
61. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2020. Available online: <https://www.R-project.org/> (accessed on 13 May 2020).
62. Van Buuren, S.; Groothuis-Oudshoorn, K. Mice: Multivariate imputation by chained equations in R. *J. Stat. Softw.* **2011**, *45*, 1–6. [[CrossRef](#)]
63. Azur, M.J.; Stuart, E.A.; Frangakis, C.; Leaf, P.J. Multiple imputation by chained equations: What is it and how does it work? *Int. J. Methods Psychiatr. Res.* **2011**, *20*, 40–49. [[CrossRef](#)] [[PubMed](#)]
64. Bürkner, P.-C. brms: An R package for bayesian multilevel models using Stan. *J. Stat. Softw.* **2017**, *80*, 1–28. [[CrossRef](#)]
65. Schneider, J.W. NHST is still logically flawed. *Scientometrics* **2018**, *115*, 627–635. [[CrossRef](#)]
66. Golino, H.F.; Demetriou, A. Estimating the dimensionality of intelligence like data using Exploratory Graph Analysis. *Intelligence* **2017**, *62*, 54–70. [[CrossRef](#)]
67. Golino, H.F.; Epskamp, S. Exploratory graph analysis: A new approach for estimating the number of dimensions in psychological research. *PLoS ONE* **2017**, *12*, e0174035. [[CrossRef](#)]

68. Golino, H.F.; Christensen, A.P.; Moulder, R. EGAnet: Exploratory Graph Analysis—A Framework for Estimating the Number of Dimensions in Multivariate Data Using Network Psychometrics. R Package Version 0.9.1. 2020. Available online: <https://CRAN.R-project.org/package=EGAnet> (accessed on 13 May 2020).
69. Jamshidian, M.; Jalal, S.J. Tests of homoscedasticity, normality, and missing completely at random for incomplete multivariate data. *Psychometrika* **2010**, *75*, 649–674. [[CrossRef](#)]
70. Jamshidian, M.; Jalal, S.J.; Jansen, C. MissMech: An R package for testing homoscedasticity, multivariate normality, and missing completely at random (MCAR). *J. Stat. Softw.* **2014**, *56*, 1–31. [[CrossRef](#)]
71. Campbell, D.T.; Fiske, D.W. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychol. Bull.* **1959**, *56*, 81–105. [[CrossRef](#)]
72. Jaussi, K.S.; Randel, A.E. Where to look? Creative self-efficacy, knowledge retrieval, and incremental and radical creativity. *Creat. Res. J.* **2014**, *26*, 400–410. [[CrossRef](#)]
73. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*; Lawrence Erlbaum Associates, Inc.: Hillsdale, NJ, USA, 1988.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).