

Reach for the stars: disentangling quantity and quality of inventors' productivity in a multifaceted latent variable model

Original

Reach for the stars: disentangling quantity and quality of inventors' productivity in a multifaceted latent variable model / Caviggioli, Federico; Forthmann, Boris. - In: SCIENTOMETRICS. - ISSN 0138-9130. - (2022). [10.1007/s11192-022-04328-3]

Availability:

This version is available at: 11583/2959781 since: 2022-03-28T18:17:27Z

Publisher:

Springer

Published

DOI:10.1007/s11192-022-04328-3

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright

(Article begins on next page)



Reach for the stars: disentangling quantity and quality of inventors' productivity in a multifaceted latent variable model

Federico Caviggioli¹ · Boris Forthmann²

Received: 25 October 2021 / Accepted: 22 February 2022
© The Author(s) 2022

Abstract

Star inventors generate superior innovation outcomes. Their capacity to invent high-quality patents might be decisive beyond mere productivity. However, the relationship between quantitative and qualitative dimensions has not been exhaustively investigated. The equal odds baseline (EOB) framework can explicitly model this relationship. This work combines a theoretical model for creative production with recent calls in the patentometrics literature for multifaceted measurement of the ability to create high-quality patents. The EOB is extended and analyzed through structural equation modeling. Specifically, we compared a multifaceted EOB model with a single latent variable for quality, and a two-dimensional model that distinguishes between technological complexity and value of invention portfolios. The two-dimensional model had better fit but weaker factor scores (for the “value” latent variable) than the unidimensional model. These findings suggest that both the uni- and the two-dimensional approaches can be directly used for extending research on star inventors, while for practical high-stakes assessments the two-dimensional model would require further improvements.

Keywords Intellectual productivity · Creativity · Patent quality · Star inventors

Introduction

Star inventors are considered of extreme interest since they generate superior innovation outcomes (Groysberg & Lee, 2009; Oldroyd & Morris, 2012; Zucker & Darby, 1997). However, their identification is not immediate and can rely on different measures: is quantity of output sufficient or are quality indicators needed? For example, stars (please

✉ Federico Caviggioli
federico.caviggioli@polito.it

Boris Forthmann
boris.forthmann@uni-muenster.de

¹ Department of Management and Production Engineering (DIGEP), Politecnico di Torino, Corso Duca degli Abruzzi, 24, 10129 Torino, Italy

² Institute of Psychology in Education, University of Münster, Fliednerstr. 21, 48149 Münster, Germany

note that from here onward we will use the terms *star inventors* and *stars* synonymously throughout the manuscript) have been usually identified through statistical cut-offs (e.g., using a top percentile or standard deviations as reference thresholds) calculated on mere productivity or on both productivity and citation counts as quality indicator (Bergé et al., 2018; Hess & Rothaermel, 2011; Rothaermel & Hess, 2007).

Disentangling inventors' productivity and the quality of their patents¹ is generally quite challenging because quantity and quality are intricately related (Forthmann, Szardenings, et al., 2021; Forthmann, et al., 2021a; Prathap, 2018; Simonton, 2009). In fact, the number of citations an inventor receives has been found to be a linear function of the number of patents, as suggested by Simonton's equal odds baseline (EOB) (Simonton, 1988a, 1988b, 1988c, 2004, 2010). However, the EOB framework also suggests the lack of a significant correlation between the average quality of an inventor's patent portfolio and its size. Recent theoretical extensions and empirical evidence strongly suggest that the quality-quantity relationship increases when conditional quantiles towards the upper tail of the quality distribution are modeled as a function of quantity (Forthmann, Leveling, et al., 2020; Forthmann, Szardenings, et al., 2021). Hence, it is questionable in how far quality indicators such as citation counts are incrementally informative for the identification of star inventors.

In addition, researchers have called for more multifaceted measurement of patent quality (Caviggioli, Colombelli, et al., 2020; Caviggioli, De Marco, et al., 2020; Higham et al., 2020; Lanjouw & Schankerman, 2004; van Zeebroeck, 2011). Consequently, a multidimensional measurement perspective that focuses on between-inventor differences and explicitly takes the quality-quantity relationship into account poses a challenge for the identification of star inventors.

In this study, we focus on patent inventors and their capacity to invent high-quality patents, a dimension of analysis which might be decisive beyond mere productivity (Kehoe et al., 2018; Rothaermel & Hess, 2007). The most frequent methods to define a star inventor in the literature considered when the inventor is either extremely prolific (quantity) and/or is involved in the creation of outstanding inventions (quality). The relationship between the two dimensions has not been exhaustively investigated in light of what can be considered determinant for the identification of star inventors as agents to increase the production of valuable innovations. We use and extend the EOB framework to explicitly take the relationship between quantity and quality into account and thus provide useful diagnostic information for the identification of star inventors. We contribute to the literature by modeling a multifaceted extension of the EOB and incorporating latent variables with the aim to isolate the measurement of quality from quantity. Hence, it should further be noted that, beyond the specific implications for the identification of star inventors, the modeling of quantity and quality focused solely on one quality dimension thus far (Den Hartigh et al., 2016; Simonton, 2004, 2010; Sinatra et al., 2016): we extend it from a unidimensional to a multidimensional quality modeling.

The remainder of this article is organized as follows. Section 2 reviews the literature on the identification of star individuals and the EOB model, deriving the aim of our research. Section 3 describes the dataset and provide details on the operationalization of the employed measures. Section 4 reports the results of the analysis. Finally, Sect. 5 concludes and discusses the findings.

¹ Notably, the perhaps most widely known attempt to combine productivity and quality in single score—i.e., the Hirsch index (Hirsch, 2005)—has also been applied to patents (e.g., Guan & Gao 2009; Motiwala et al., 2020).

Research framework

Star individuals

The relevance of star scientists is not limited to a direct increase of output (Groysberg & Lee, 2009) but they also support organization activities (Kehoe & Tzabbar, 2015) and improve the attraction of resources and skilled personnel (Hess & Rothaermel, 2011; Lacetera et al., 2004). They also indirectly foster the productivity of peers and collaborators thanks to learning and emulation (Lockwood & Kunda, 1997). Although there is consensus on the presence of a general positive impact of star individuals, it is worth reminding that in some cases the literature identified negative effects in organizations due to coordination costs and conflicts (Bendersky & Hays, 2012; Groysberg et al., 2011; Swaab et al., 2014). Furthermore, hiring stars is often expensive (Groysberg et al., 2011) and thus it should be considered a critical activity. The findings of the literature support the need to improve the understanding of the way to identify exceptional scientists.

The identification of stars has taken different approaches in the literature, with respect to the examined field of activity and the different operationalizations of the criteria to distinguish outstanding from common individuals. In general, to be a star the individual must engage in disproportionately high performance relative to most other workers in their field (Aguinis & O'Boyle, 2014; Call et al., 2015). The examined performance has been measured under different perspectives ranging from productivity (Kehoe & Tzabbar, 2015; Lahiri et al., 2019; Subramanian et al., 2013; Zucker et al., 2002), impact (Azoulay et al., 2010; Rothaermel & Hess, 2007) and, in some cases, visibility or celebrity (Oldroyd & Morris, 2012).

Star individuals have been studied in several contexts² with particular attention to scientists/scholars (Azoulay et al., 2010) and inventors (Hohberger, 2016), thanks to data availability, i.e. articles and patents. Stars in these two categories have been similarly addressed by considering either their productivity in terms of quantity of output, in most cases through the number of articles or patents, their impact relying on a measure of quality such as the received citations (Hess & Rothaermel, 2011; Hohberger, 2016; Liu, 2014), or a combination of them (Agrawal et al., 2017; Kehoe & Tzabbar, 2015). The extent to which performance must be disproportional varies across studies (Call et al., 2015): some have used from one to three standard deviation (SD) difference (e.g. Hess & Rothaermel, 2011), others have applied a cutoff value for the top percentage of the examined sample, from 1 to 10% (e.g. Hohberger, 2016). Table 1 summarizes the approaches in the literature.

So far the literature dealing with the identification and the analysis of the role of star inventors has not considered that patents can be evaluated for a variety of quality criteria and researchers have called for a multifaceted perspective on patent quality (e.g. Lanjouw & Schankerman, 2004; van Zeebroeck, 2011). In particular, the measurement of patent quality can be decomposed in two main dimensions with respect to the nature of the protected invention: technological complexity and value (Caviggioli, Colombelli, et al., 2020; Caviggioli, De Marco, et al., 2020; van Zeebroeck & van Pottelsberghe, 2011). Technological complexity refers to the number of components, their degree of inter-dependence and decomposability (Singh, 1995; Wang et al., 2013). Patent value, as conceptualized in this work, refers to the technical merit and the potential market size of the invention

² For example, sport players (Chen and Garg 2018) and actors (Han and Ravid 2020).

Table 1 Methods for the identification of stars in previous works

Study	Quantity	Quality	Additional criteria	Percentage of stars in the sample
(Zucker & Darby, 1997)	Genetic sequence discoveries		Three-year moving average of annual publications > 5 for at least 1 year	Not calculable or reported
(Lacetera et al., 2004)	Publications on genetic sequence discoveries		Three-year moving average of annual publications > 5 for at least 1 year	2.58%
(Rothaermel & Hess, 2007)	Publications	Citations to publications	Publication count > 3SD above the mean	1.77%
	Publications	Citations	Citation count > 3SD above the mean	1.16%
		Nobel Laureates	Publication count > 3SD above the mean AND Citation count > 3SD above the mean	0.63%
(Azoulay et al., 2010)			At least one of the following criteria: (1) highly funded scientists (> 95th p.); (2) highly cited scientists (top 250); (3) top percentile patenters; and (4) members of the National Academy of Sciences (5) NIH-MERT awardees; (6) Howard Hughes medical investigators; and (7) early career prize winners	0.02%
				5%
(Hess & Rothaermel, 2011)	Publications	Citations	Publication count > 3SD above the mean AND Citation count > 3SD above the mean	0.63%
(Beaudry & Schiffrerova, 2011)	Patents		Patent count > 20	Not calculable or reported
(Kehoe & Tzabbar, 2015)	Patents	Citations	Productivity score as yearly number of patents weighted by received citations and compared to industry average. Star if score > 2SD above the mean	7%
(Hohberger, 2016)		Patent citations	Cumulative patent forward citations per inventor; one percent cut-off criterion	1%

Table 1 (continued)

Study	Quantity	Quality	Additional criteria	Percentage of stars in the sample
(Agrawal et al., 2017)	Publications	Citations	Scientist whose stock of citation-weighted articles published up until year $t - 1$ is above the 90th percentile (or 95th)	10% (5%)
(Bergé et al., 2018)	Patents		Patent count $> Q_{99}$ (Patent count)	Not calculable or reported

Notes. The examined samples are different across the studies

(Caviggioli, Colombelli, et al., 2020).³ The corresponding measures will be described in detail in Sect. 3.2.

Quantity and quality: the equal odds baseline

Previous literature analyzed the relationship between quantitative and qualitative output of intellectual activities and reported mixed evidence. According to one approach of the literature dealing with creativity, high-quality ideas consume time and resources, either intellectual and physical: this suggests the presence of a trade-off and a negative correlation between average quality (per product) and quantity (Fischer et al., 2012; Guilford, 1968; Michalska-Smith & Allesina, 2017). On the other hand, the dual pathway of creativity (Nijstad et al., 2010) considers a positive correlation between average quality and quantity, achieved through two behaviors: flexibility, in terms of variety of conceptual ideas, or persistence-and-exhaustion in terms of specialization on a focal theme. Yet, other models emphasize the role of luck and propose a null correlation between quantity and average quality (Janosov et al., 2020; D. K. Simonton, 2010; D K Simonton, 1988a, 1988b, 1988c; Sinatra et al., 2016). In this work, we focus on the EOB which belongs to the latter group of models.

Extending the previous work by Wayne Dennis (1958), the seminal study of (Simonton, 1988a, 1988b, 1988c) introduced the EOB, a statistical model for the relationship between quantity and quality of scientific output within a comprehensive theoretical framework for scientific productivity (Simonton, 2009, 2010). Considering the focus of this study on patents, the EOB relies on two main propositions. First, the number of an inventor' high-quality patents H (i.e., the number of hits) is positively and linearly related to the total number of patents T . Previous works (e.g. Forthmann, Leveling, et al., 2020; Forthmann, Szardenings, et al., 2021) employed the number of citations received as a measure to identify "hits". The EOB models the following equation (Simonton, 2010; p. 163):

$$H_i = \rho T_i + u_i. \quad (1)$$

where ρ refers to the hit-ratio and u_i is a random error term for inventor i .⁴ The second proposition we highlight in the EOB framework is that individual hit-ratios H/T are uncorrelated with T (i.e., invention portfolio sizes in the context of this study). Otherwise it would follow that the relationship between H and T is non-linear (Simonton, 2003, 2004). In other words, a positive linear correlation between H and T is a necessary but not sufficient condition for the EOB (Forthmann et al., 2021b; Forthmann, Szardenings, et al., 2021; Forthmann, Leveling, et al., 2021). The EOB further proposes an intercept of zero and a hit-ratio that equals the ratio of average H and average T . These implications of the EOB allow evaluation of model fit within the framework of structural equation modeling, SEM (Forthmann et al., 2020; Forthmann et al., 2021). SEM is a widely used approach in sociological and psychological research, for example, and it provides many options for the evaluation of data-model fit (West et al., 2012).

³ Ideally, one would want to assess value also in terms of economic relevance of the protected invention but unfortunately this type of information cannot be easily determined for large samples. The amount of money a patent is worth can be determined when it is licensed or sold in disclosed transactions (e.g., cases when data on auctions are revealed, as in Caviggioli & Ughetto, 2016), or as a means of broad surveys (e.g., Torrisi et al., 2016).

⁴ Please note that we omit the subscript i for simplicity in many parts of the paper.

A multifaceted extension of the equal odds baseline

As mentioned above, researchers have called for a multifaceted perspective on patent quality (Lanjouw & Schankerman, 2004; van Zeebroeck, 2011). Previous literature on the EOB, however, has not yet considered multiple dimensions of quality within the same modeling framework (nor did other chance models of scientific productivity (see, for example, Janosov et al., 2020 and Sinatra et al., 2016)). Hence, in this work we rely on the empirical framework of patent inventors that makes it possible to leverage the presence of several indices of quality. This comprehensive assessment approach provides support to disentangle quantity and quality, with quality as a multifaceted dimension. Notably, beyond the concrete aim of constructing an assessment model for inventors' capacity to create high-quality patents, our work extends chance models of creative success from unidimensional models to multidimensional modeling of quality. We argue that this approach allows a direct test of the generalizability of the model that can be evaluated in a multivariate model (i.e., not in multiple analyses conducted separately).

Specifically, this work aims at extending the EOB in this regard by formulating the EOB as a SEM in which individual differences in hit-rates are explained by a quality latent variable. In other words, the error term u_{ij} for inventor i ($i = 1, \dots, I$) and quality indicator j ($j = 1, \dots, J$) will be modeled by the following equation

$$u_{ij} = \lambda_j \eta_i + \varepsilon_{ij}, \tag{2}$$

with latent factor η_i as the capacity to create high-quality inventions, λ_j being the loading of the j th indicator on the capacity factor, and ε_{ij} the remaining error left unexplained after taking quantity and the capacity for quality into account. Inserting Eq. 2 into Eq. 1 yields a multifaceted EOB:

$$H_{ij} = \rho_j T_i + \lambda_j \eta_i + \varepsilon_{ij}. \tag{3}$$

Specifically, the multifaceted EOB proposes that η and T are uncorrelated which allows for independent assessment of inventors' capacity for quality and productivity (i.e., quantity of output). As an extension of Eq. 1, the model in Eq. 3 can also be estimated within the SEM framework (Bollen, 1989). In SEM, a proposed path model, its implied covariance matrix and mean vector are examined for their discrepancy to their empirically observed counterparts. Useful models have a model-implied covariance matrix and mean vector that are close to the observed covariance matrix and vector of means (Bollen, 1989). Goodness of fit between a proposed model and data in SEM can be evaluated by various established indices (West et al., 2012). In this approach, regression coefficients ρ_j and λ_j can be estimated by maximum likelihood (or other robust variants). Estimates of the latent capacity η_i can be obtained by means of empirical Bayes (Estabrook & Neale, 2013), for example. This model is illustrated for the five quality indicators used in this study on the left side in Fig. 1 (further details in Sect. 3). In particular, two models will be tested: one where capacity for quality is unidimensional (Model 1 in Fig. 1) and a second model where two latent variables for quality are assumed (Model 2 in Fig. 1). Since the measurement of patent quality can be decomposed into the dimensions of technological complexity and value (Caviggioli, Colombelli, et al., 2020; Caviggioli, De Marco, et al., 2020; van Zeebroeck & van Pottelsberghe, 2011), Eq. 3 needs to be extended to a two-dimensional model that includes the two corresponding latent variables (see also Model 2 in Fig. 1 for a path model illustration).

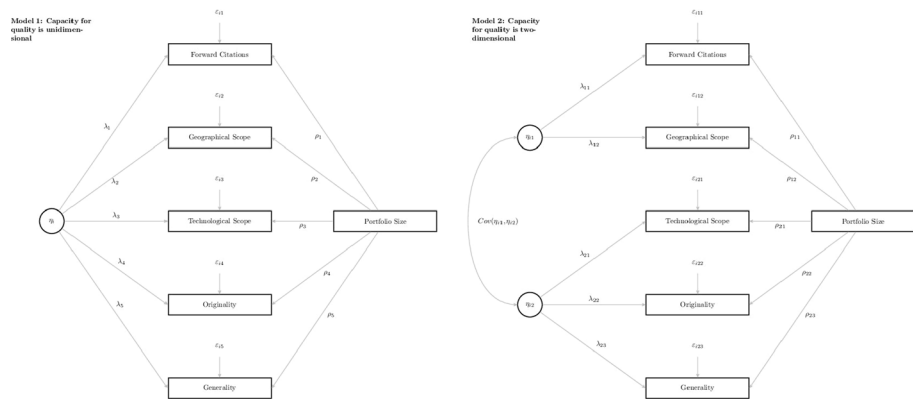


Fig. 1 Path models of the unidimensional (left) and two-dimensional (right) multifaceted EOB models illustrated with the five quality indicators used in this study (see Sect. 3.2 for details)

We aim to empirically test the fit of data on inventors to the EOB when quality is measured by multiple indicators and to explain the hit-ratio variation as a means of a latent capacity to create high-quality inventions. This study extends the recent findings and theorizing on the EOB in several ways. First, the literature on EOB has so far focused on quality measured with forward citations of patents (Forthmann et al., 2020, 2021b): other indicators are introduced because patent quality is a multifaceted construct (Lanjouw & Schankerman, 2004; van Zeebroeck & van Pottelsberghe, 2011). Second, the EOB is extended beyond the current results (Forthmann et al., 2021) and in this work it incorporates a latent quality variable that potentially explains individual differences in hit-ratios within a SEM framework. Within the classical approach to EOB there is only one quality score H and one quantity score T and the differences in hit-ratios are explained by the residual term (e.g., some researchers are luckier than others). The residual term also reflects quality as differences in hit-ratios, but with only one quality score H it is not possible to isolate individual differences in the hit-ratio. In a multifaceted approach, there are as many residual terms as quality indicators and latent quality factors can be measured based on these residuals. This approach makes it possible to measure quality independent of quantity, while at the same time H depends linearly on T .

Aim of the current research

The main aim of this study is to extend recent findings and theorizing on the EOB in several ways. First, we extend recent results obtained for forward citations of patents (Forthmann et al., 2021b; Forthmann, Leveling et al., 2020; Forthmann et al., 2021) to other indicators because patent quality is a multifaceted construct (Lanjouw & Schankerman, 2004). Second, new EOB theorizing allows quantifying the amount of residual variance accounted by mere sampling variation (Forthmann et al., 2020; Forthmann et al., 2021). This is useful to accurately estimate the amount of hit-ratio variation that is attributable to between-inventor differences. The presence of between-inventor variation in hit-ratios is essential for the measurement of capacity for patent quality. In this vein, the EOB is extended in this work to incorporate a latent quality variable that potentially explains individual differences in hit-ratios within a SEM framework. Finally, we aimed at comparing

a unidimensional model with a two-dimensional model that incorporated latent variables to measure both technological complexity and patent value. Importantly, a reasonable fit of the data to either the unidimensional or two-dimensional multifaceted EOB implies that inventors' capacities to create quality patents can be measured to potentially identify star inventors in a way that explicitly takes the intricate relationship between overall productivity and patent quality into account. Finally, the feasibility of the outlined approach for practical assessment contexts (e.g., high-stakes decisions) was examined. The approach results in estimates of inventors' capacity to invent high-quality patents. These estimates taken from the multifaceted EOB are factor scores and it has been recommended in the literature that the correlation between these estimates and their true values (i.e., the factor determinacy index; FDI) should be larger than 0.80 for research purposes, but larger than 0.90 for individual differences assessments in practical high-stakes contexts (Ferrando & Lorenzo-Seva, 2018). Thus, we wanted to examine if inventors' capacity estimates can achieve this level of quality and further aimed to illustrate their usage for the identification of stars in comparison with other commonly applied approaches.

Method

Data sources

The main data sources are PatentsView and PATSTAT. PatentsView is a data warehouse sourced from USPTO-provided data on published patent applications (2001-present) and granted patents (1976-present). It provides disambiguated inventors' names from the application of an algorithm that uses discriminative hierarchical co-reference.⁵ Patent level data from PatentsView are linked to PATSTAT, the largest repository of patent data in terms of coverage and available information, maintained by the EPO with the collaboration of the main patent offices.⁶

The analyses will be carried out at the level of inventors and the examined sample is defined by applying the following steps. The starting sample includes all the inventors with at least one US granted patent filed between 2008 and 2010,⁷ corresponding to 725,577 disambiguated names in PatentsView. The selected inventors are associated to a total of 4,297,710 granted patents (their full patenting history) which are linked to PATSTAT to collect further information.⁸

All the selected patents are associated to their INPADOC family (2.9 million families). Patent families represent a unit of analysis that is closer to invention: multiple patent documents regarding the same filing are collapsed to a single unit, providing a more accurate measure of inventors' productivity (OECD, 2009; Martínez, 2011). Furthermore, country extensions can be identified providing information on the geographical coverage. The

⁵ More information at www.patentsview.org (last access and data release in August 2020, disambiguated inventors' names updated in March 2020).

⁶ More information at <https://www.epo.org/searching-for-patents/business/patstat.html> (last access in August 2020, version of database used in this study: fall 2019).

⁷ Only "utility" patents have been considered. Withdrawn patents are included (corresponding to 0.17% of the examined granted patents).

⁸ In terms of granted patents, the selected sample is 58% of the total US granted patents in PatentsView (1975–2020).

Table 2 Distribution of portfolios of inventions across selected inventors

Portfolio size [from	to]	Number of inventors	Perc.	Cumulate
1	1	214,226	30.43	30.43
2	5	252,228	35.83	66.26
6	10	108,368	15.40	81.65
11	50	119,041	16.89	98.56
51	100	8,032	1.14	99.70
101	Max	2,082	0.30	100.00
	Total	703,977	100%	

earliest filing year and the IPC subclasses of each family are collected and several patentometrics are calculated following the approach described in (Caviggioli, Colombelli, et al., 2020; Caviggioli et al., 2020). For each inventor it is thus possible to identify the portfolio of inventions and create portfolio level measures (these variables are described in the next section in detail) as of 2010 in terms of productivity. The cut-off year is required to consider a subsequent time window sufficiently large to calculate quality indicators such as citations and to account for potential delays in the publication of documents.

With the aim to clean the sample from potential errors in the original data, either in name disambiguation or in patent family identification, those inventors reporting a portfolio-level earliest filing date prior to 1981 (3.2%) were excluded. Inventors with no IPC codes associated to the invention portfolio were also eliminated (0.01%).

The final sample is a selection of 703,977 inventors active in the years 2008–2010 and with a patenting history of maximum 30 years in 2010: each invention portfolios represent the cumulated inventions up to 2010. Table 2 reports the distribution of the portfolio size in the sample.

Variables

The inventors' quantitative productivity is captured by the number of patent families between 1981 and 2010, corresponding to the T in the previous formulas. Quality is described through the count of an inventor's outstanding inventions H according to several measures with the aim to test their relationship with quantity.

The first step to generate the indicators of quality was to compute for any patent family the corresponding value of technological scope, generality and originality index, forward citations and geographical scope. The first three refer to indicators of technological complexity and the last two to the value of an invention.

The technological scope counts the number of different IPC subclasses associated to patents (Lerner, 1994): the count is extended to the family level by considering all the family members. It provides a measure of multi-disciplinarity: the broader the scope, the greater the complexity and the potential range of technological areas where it can impact (Harhoff et al., 2003).

The originality and generality indexes are variations from the Simpson diversity index, also known as the Hirschmann-Herfindahl index or the repeat rate (for diversity; e.g., Rousseau, 2018). They were first introduced in the patent data framework by Trajtenberg et al. (1997) and are calculated considering the concentration of the different technological fields among the cited and citing patents of every focal document respectively:

$$\text{Generality}_f = 1 - \sum_{k=1}^{N_f} \left(\frac{\#\text{Citing Patents}_{fk}}{\#\text{Citing Patents}_f} \right)^2 \tag{4}$$

$$\text{Originality}_f = 1 - \sum_{k=1}^{N_f} \left(\frac{\#\text{Cited Patents}_{fk}}{\#\text{Cited Patents}_f} \right)^2 \tag{5}$$

Any patent family f is associated to k technological fields (up to N_f fields), identified by IPC subclasses (four-digit IPC codes). Coherently with the general approach, the patent citation network is generated at the level of INPADOC families, excluding intra-family citations. The generality index is a forward-looking measure describing the width of the technological advances. The originality index represents the scope of the underlying research.

The forward citations provides a measure of the technical value (van Zeebroeck and van Pottelsberghe, 2011). The indicator considers only citations occurring in the first five years after the filing to account for the different time of exposure to the “risk” of receiving a citation (Caviggioli & Ughetto, 2016).⁹

The geographical scope indicates how large the expected market for the patented technology is. It is calculated as the number of jurisdictions in which patent protection is sought (Agostini et al., 2015; Lanjouw et al., 1998).

Once each patent family was associated to its measures of quality, the next step was to follow the approach proposed in van Zeebroeck (2011) which allows calibrating the indicators with respect to the specificities of technological areas and the potential trends occurring in the time frame. Coherently with the unit of analyses, the approach was applied at the patent family level. The calibration of each indicator of quality is achieved by ranking patent families in a reference cohort, defined by technological sector and year. The sectors are identified by considering the concordance table between the IPC codes and 35 technical fields, developed by the WIPO.¹⁰ The reference time is the earliest filing year among the family members. The ranking leads to a percentile value for each patent family, ranging between 0 and 100. It represents the share of families in the same sector and with the same earliest filing year that have a lower score than the examined family.¹¹ When a patent family is associated to more technical fields, the indicator assumes the value of the highest percentile. For example, if an invention is developed in the fields “Optics” and “Pharmaceuticals” and the percentile of the examined family is 60th for forward citations among all the inventions in the former and 80th in the latter, then the selected score for the considered family is 80.

⁹ The models were also tested considering a variable with a 10 years window to capture citations (as in Forthmann et al., 2020a). The results are very similar and are openly available in the OSF repository (<https://osf.io/bjad4/>). Note that intra-family citations are not considered.

¹⁰ Source: WIPO IPC-Technology Concordance Table (last update in 2016), available at https://www.wipo.int/ipstats/en/statistics/patents/xls/ipc_technology.xlsx, last access August 2021. Further info available in Schmoch (2008).

¹¹ For example, patent families with zero forward citations report “0” as indicator (no other family has a lower number of forward citations). The patent family with the maximum number of citations in the sector-year cohort would be higher than 99.99% of the other families in the same group and thus the indicator reports a rounded value of 100.

An invention is thus considered outstanding according to the quality indicator j when it is equal or above the 95th percentile in the corresponding sector-year cohort (i.e., the family level indicator is equal or above 95):

$$\text{Outstanding Invention}_j = \begin{cases} 1, & \text{Indicator}_j \geq 95 \\ 0, & \text{Indicator}_j < 95 \end{cases} \quad (6)$$

Note that a single patent family could be above the excellence threshold in none, one or more of the indicators of quality. Once all the top inventions are identified with respect to the quality indicator j , they are aggregated at the portfolio level for each inventor i . This provides the number of outstanding inventions generated by inventor i (i.e., her/his hits H):

$$H_{ij} = \sum_{k=1}^{N_i} \text{Outstanding Invention}_{kj}. \quad (7)$$

The hit ratio, that is the share of outstanding inventions in the inventor's portfolio, can be calculated by dividing H with the portfolio size T (i.e., H/T).

Finally, the following two variables are included to improve the model specifications and control for inventor's characteristics. Since the selected sample includes inventors at different stage of their career, a proxy of their expertise is introduced as the number of years since the first filing date. PatentsView database provides also data on inventors' gender, as a result of the method explained in the report of the Office of the Chief Economist (2019). Note that data coverage is not complete (the gender is missing for 9.1% of the inventors in the examined sample).

Table 3 shows summary statistics of the variables.

Analysis

Initially, we examined the fit of the data to the EOB for each of the five quality indicators by means of correlations and a check of the presence of individual differences in the residual (Forthmann et al., 2020; Forthmann et al., 2021). In the next step, we fitted the multifaceted EOB in SEM framework. Finally, we examined star identification based on the multifaceted EOB as compared to other common approaches for star identification (see Table 1). The R script used to perform the reported analyses and an html-file with all the related output are openly available in an online repository in the Open Science Framework (<https://osf.io/bjad4/>).

Preliminary tests

First, positive correlations between patent family counts (T) and all indicators of quality (H) were found (see column 1 in Table 4). This is in accordance with the EOB which proposes that the relationship between H and T is positive and linear. Concerning the control variables, as expected, career length was moderately positively correlated with family count and with small/moderate values with all indicators of quality. Gender did not correlate with any of the measures for creative productivity.

Second, the correlations between patent family counts and each of the quality indicators expressed as hit-ratio (H/T) were close to zero (see column 1 in Table 5), which again provides evidence in favor of the EOB across all quality indicators. Both the

Table 3 Summary statistics of the examined variables

Measure	Indicator	Obs.	Mean	Std.Dev	Median	Min.	Max.
<i>Quantity (T)</i>							
Inventor's portfolio size	Count of inventions as patent families having earliest filing year in 1981–2010	703,977	7.25	13.82	3	1	1041
<i>Quality of the portfolio of inventions as</i>							
<i>Count of patent families above the 95th percentile for the following indicators (H):</i>							
Tech. complexity–Multidisciplinarity	Technological scope	703,977	0.26	0.87	0	0	103
Tech. complexity–Redeployability	Generality index	703,977	0.50	1.60	0	0	176
Tech. complexity–Novelty	Originality index	703,977	0.49	1.68	0	0	294
Value–Technical merit	Count of Fwd cit. (5-years window)	703,977	0.59	1.93	0	0	198
Value–Potential market size	Geographical scope	703,977	0.44	1.74	0	0	197
<i>Quality of the portfolio of inventions as</i>							
<i>Share of patent families above the 95th percentile for the following indicators (HT):</i>							
Tech. complexity–Multidisciplinarity	Technological scope	703,977	0.05	0.16	0.00	0	1
Tech. complexity–Redeployability	Generality index	703,977	0.07	0.20	0.00	0	1
Tech. complexity–Novelty	Originality index	703,977	0.07	0.20	0.00	0	1
Value – Technical merit	Count of Fwd cit. (5-years window)	703,977	0.08	0.20	0.00	0	1
Value–Potential market size	Geographical scope	703,977	0.07	0.20	0.00	0	1
<i>Control variables</i>							
Expertise – career	Years from the earliest filing date	703,977	8.00	6.76	5	1	30
Gender	Dummy = 1 if male	640,043	0.87	0.34	1	0	1

Table 4 Correlation matrix of examined variables: quality indicators expressed as number of excellent inventions in the portfolio (Testing $H \sim T$ —EOB correlations between quantity and quality indicators in bold)

Variables	1	2	3	4	5	6	7
1 Quantity: Count of families	1.00						
2 Quality: Tech. scope	0.43	1.00					
3 Quality: Generality index	0.49	0.65	1.00				
4 Quality: Originality index	0.47	0.61	0.78	1.00			
5 Quality: Count of fwd.cit. (5-yrs)	0.61	0.47	0.46	0.42	1.00		
6 Quality: Geographical scope	0.39	0.33	0.25	0.22	0.41	1.00	
7 Career	0.38	0.24	0.29	0.26	0.24	0.18	1.00
8 Gender	0.05	0.03	0.04	0.03	0.03	0.02	0.11

Table 5 Correlation matrix of examined variables: quality indicators expressed as ratio of excellent inventions on total number of inventions in the portfolio (Testing $H/T \sim T$ —EOB correlations between quantity and quality indicators in bold)

Variables	1	2	3	4	5	6	7
1 Quantity: Count of families	1.00						
2 Quality: Tech. scope	-0.03	1.00					
3 Quality: Generality index	-0.02	0.33	1.00				
4 Quality: Originality index	-0.02	0.33	0.39	1.00			
5 Quality: Count of fwd.cit. (5-yrs)	0.00	0.32	0.21	0.20	1.00		
6 Quality: Geographical scope	-0.03	0.23	0.09	0.09	0.19	1.00	
7 Career	0.38	0.04	0.08	0.07	0.08	0.03	1.00
8 Gender	0.05	-0.01	0.01	0.00	0.01	-0.01	0.11

control variables, career length and gender, correlated with all the variables of quality in terms of hit-ratio in a negligible way.

As a final preparatory step prior to the examination of the multifaceted EOB, we checked if residual variances were larger than what is expected under the strict EOB, the model which implies a constant hit-ratio, or in other words where the error term u is excluded. This check can be meaningfully done when the EOB displays reasonable fit, as suggested by the correlations reported in Tables 4 and 5. The SEM approach to study the EOB has been recently further extended (Forthmann et al., 2021b) to allow quantifying if residual variance $Var(u)$ is larger as compared to the strict EOB with $Var(u) = 0$. This approach can be used to examine if individual differences are present in a given dataset, i.e., hit-ratio variance is larger than mere sampling error variation. The presence of individual differences in hit-ratios are a prerequisite to measure quality as a latent variable based on residuals resulting from multiple quality indicators.

Residual variance findings are reported in Table 6. All observed residual variances were at least twice as large as compared to the minimum expected residual variance under the EOB (i.e., the residual variance under strict equal odds). Hence, we conclude that hit-ratio variation in the data was larger than expected under strict equal odds which

Table 6 Estimates of the observed residual variance and the smallest expected residual variance

Variables	Observed residual variance [95% CI]	Smallest residual variance (Strict EOB) [95% CI]
Quality: Tech. scope	0.608 [0.577, 0.639]	0.255 [0.253, 0.257]
Quality: Generality index	1.937 [1.820, 2.054]	0.463 [0.460, 0.466]
Quality: Originality index	2.186 [1.938, 2.434]	0.455 [0.451, 0.458]
Quality: Count of fwd.cit. (5-yrs)	2.353 [2.21, 2.49]	0.540 [0.536, 0.544]
Quality: Geographical scope	2.555 [2.360, 2.750]	0.411 [0.408, 0.415]

only allows sampling error as a unique source of residual variation. Hence the data were promising for the application of the multifaceted EOB with latent variable(s).

Multifaceted EOB results

The SEM framework is implemented with the package *lavaan* (Rossee, 2012) for the statistical software R. Model fit was based on indices such as the RMSEA, SRMR, CFI, and TLI (Table 7), according to existing cut-offs in the SEM literature (West et al., 2012). Using these fit indices is particularly helpful when examining the EOB in very large datasets because even small and negligible deviations from the EOB become easily statistically significant (Forthmann et al., 2020; Forthmann et al., 2021). SEM model fit indices indicate if the data can be adequately described by the EOB when sample sizes are large. Finally, we estimated marginal reliability for the latent quality variables to quantify measurement precision (Brown & Croudace, 2015; Green et al., 1984). We further report the factor determinacy index (FDI; i.e., the correlation between estimated factor scores and their true values) which can be obtained as the square-root from marginal reliability. The FDI is a useful index to quantify the measurement quality of factor scores for subsequent assessment purposes. For example, it has been proposed that a $FDI > 0.80$ is sufficient for research purposes, whereas a value for the $FDI > 0.90$ is needed for the assessment of individual differences in high-stakes situations (Ferrando & Lorenzo-Seva, 2018).

The multifaceted EOB without any latent quality variables displayed excellent fit. The model with a single latent variable for quality, the unidimensional one, was estimated and displayed adequate fit, with TLI being the only index that did not pass the common cut-off of 0.95. Standardized factor loadings of this model are depicted on the left side of Fig. 2. These loadings revealed that the overall quality factor was dominated by technological complexity (standardized loadings were in the range from 0.57 to 0.76). The indicators of quality in terms of “value” had only small loadings (the variables on forward citations and geographical scope had a loading of 0.22 and 0.10 respectively). Marginal reliability of the latent quality variable was 0.86 indicating good reliability. In addition, the FDI was equal to 0.93, which indicated that factor score estimates based on the unidimensional multifaceted EOB model had excellent quality that allows using them in the context of high-stakes decisions (Ferrando & Lorenzo-Seva, 2018). The latent quality variable was predicted by career length to a small degree ($\beta = 0.14, p < 0.001$) and negligibly small by gender ($\beta = -0.02, p < 0.001$) with an overall $R^2 = 0.02$.

We outlined in the introduction that patent quality can be defined through multiple measures, each with different nuances. For the current work, we decided to stick to the

Table 7 SEM fit indices for the model with no latent variable and for the unidimensional and the two-dimensional models

Fit indices	No latent variable	Unidimensional model	Two-dimensional model
Root Mean Square Error of Approximation (RMSEA)	.004 90%–CI: [.004, .005]	.015 90%–CI: [.015, .015]	.013 90%–CI: [.013, .013]
Standardized Root Mean Square Residual (SRMR)	.007	.043	.034
Comparative Fit Index (CFI)	.999	.960	.977
Tucker Lewis index (TLI)	.999	.939	.962

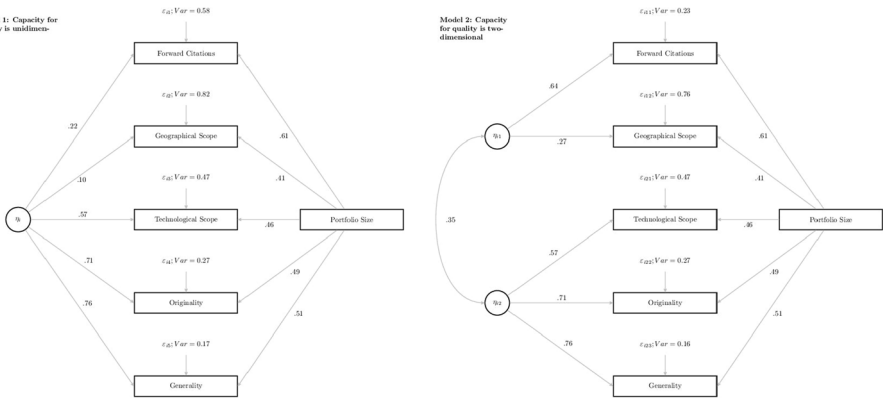


Fig. 2 Standardized estimates for the unidimensional (left) and two-dimensional (right) multifaceted EOB models

dimensions of technological value and complexity (e.g., Caviggioli, Colombelli, et al., 2020; van Zeebroeck and van Pottelsberghe de la Potterie 2011). Notably, those studies have included the generality index among the measures of complexity, giving more importance to the aspect of embedding multiple features for several applications rather than a higher number of citations per se, which is directly addressed by the count of forward citations. However, generality as an index that is also based on forward citations could have been alternatively proposed to load on the value factor rather than technological complexity. Hence, we tested two alternative models: (a) a model in which generality loaded on value and not on complexity, and (b) a model with cross-loadings of generality on both latent factors. The results suggest that these alternative model specifications were not better than the original ones (see Fig. 1). First, we tested a two-dimensional model in which generality loaded on value instead of complexity and found that the latent covariance matrix was not positive definite. Inspecting the matrix, it turned out that the correlation between both quality dimensions for such a model was greater than one. Such an anomalous finding can occur in SEM and it can hint at a mis-specified model, for example. Another alternative model might allow cross-loadings of generality on both value and complexity. Such a model was estimated without any technical difficulties. However, it also did not outperform the intended two-dimensional model depicted in Fig. 1 and for which results are reported below. The two-dimensional model based on previous literature (RMSEA=0.013, SRMR=0.034, CFI=0.977, TLI=0.962) had mostly better fit indices as compared to the model with cross-loadings for generality (RMSEA=0.015, SRMR=0.034, CFI=0.978, TLI=0.958). In addition, the standardized loading of generality on the value factor was -0.03 and, hence, negligibly small and negative. Consequently, we consider these additional checks as further validity evidence in favor of our initially intended models.

With respect to the unidimensional model, the two-dimensional model with latent variables for value and technological complexity displayed better fit. The factors that describe technological value and complexity factors reported a correlation of 0.35, indicating a limited overlap. Figure 2 reports the standardized factor loadings (right side): the indices of technological complexity display strong factor loadings (all above 0.57), while for value, only the variable based on forward citations shows a strong loading (0.64). In this model, marginal reliability was 0.66 for the latent variable referring to the dimension of “value”

and 0.86 for the latent variable referring to the technological complexity. The factor scores for “value” had an FDI equal to 0.81 and can thus be used for research purposes, but not for the practical assessment of individual differences (Ferrando & Lorenzo-Seva, 2018). The factor scores for the technological complexity were associated to an FDI of 0.93, thus having excellent psychometric quality: they can be used for any assessment purpose (e.g., high-stakes decisions). Hence, value was comparably less reliable, whereas technological complexity had excellent reliability. The results suggest that the measurement of value should be complemented by other indicators developed in the corresponding samples: for example the number of renewals of granted patents, litigations or oppositions for disputed patents, licensing or sales data for transacted inventions, or direct assessment of the economic relevance through a survey (Cavaggioli & Ughetto, 2016; Torrisi et al., 2016; van Zeebroeck and van Pottelsberghe 2011).

The technological complexity latent variable was predicted by career length to a small degree ($\beta=0.14$, $p<0.001$) and negligible small by gender ($\beta=-0.02$, $p<0.001$) with an overall $R^2=0.02$, similarly to the findings for the unidimensional model. In addition, the R^2 for the “value” latent variable was zero which indicated that both control variables had a negligible relationship with it.

To test the robustness of the results, we analyzed the same models on 35 subsamples based on the technological fields represented in the WIPO concordance table. Each inventor was associated to one or more fields by considering the technological areas where s/he patented the most or representing at least a third of her/his total portfolio of inventions, to avoid marginal contributions. The fit indices of the 35 models are very similar across the subsamples based on the WIPO fields. The results are reported in the Appendix (Fig. 3).

Identification of stars

To see the proposed assessment approach based on the multifaceted EOB in action, we compared estimates of inventors’ capacity to create high-quality patents with other commonly used approaches for the identification of stars. To understand how far quality indicators are incrementally informative for the identification of stars with respect to quantity alone, we calculated the Jaccard similarity between different groups of stars identified by different criteria. We tested in our sample if using either productivity or citation counts identifies nearly the same set of stars, and if so, how well the introduction of the multifaceted EOB for quality measurement can produce incremental information when the multifaceted EOB is used for measurement of quality: Jaccard similarity quantifies the amount of incremental information provided by different approaches for star identification beyond quantity.

We identified stars as inventors who performed better than $3SD$ above the mean because this strategy has been used the most in the literature (Table 1). We employed this criterion for star identification to the dimensions of quantity, of quality based on forward citations, as well as to both quantity and the number of high-quality patents based on forward citations (Table 8). All these variables were normalized for career length (to account for the correlations of Table 4). These common approaches yielded percentages of stars (between 0.32 and 1.31%) in a comparable range as in previous works (Table 1). In addition, we used factor scores from the unidimensional multifaceted EOB model (Model 1 on the left in Figs. 1 and 2) because the FDI passed the recommended cut-off for usage in high-stakes decision contexts. Hence, for practical purposes we chose a model with less favorable model fit results (yet model fit was still adequate) over a better fitting model because the

Table 8 Percentages of stars, absolute star samples, and Jaccard similarity between various approaches to select stars (M = mean, 3SD = three standard deviations)

Variables	Per. of stars [95% CI]	Number of stars	1	2	3	4
1 Quantity: Count of families > M + 3SD	1.31 [1.28, 1.33]	9195				
2 Quality: Count of fwd.cit. > M + 3SD	0.72 [0.70, 0.74]	2392	14.30 [13.63, 15.00]			
3 Quantity: Count of families > M + 3SD AND	0.32 [0.30, 0.33]	2234	24.30 [23.42, 25.19]	45.65 [43.91, 47.41]		
Quality: Count of fwd.cit. > M + 3SD						
4 Quality: η (Model 1) > M + 3SD	1.01 [0.99, 1.04]	7135	7.01 [6.61, 7.43]	7.31 [6.78, 7.87]	5.90 [5.42, 6.41]	
5 Quantity: Count of families > M + 3SD AND	0.15 [0.14, 0.16]	1070	11.64 [10.99, 12.31]	12.77 [11.61, 14.00]	18.76 [17.33, 20.26]	15.00 [14.18, 15.85]
Quality: η (Model 1) > M + 3SD						

All variables used for the identification of stars were divided by career length to control for this variable

reliability of the associated capacity estimates indicated unambiguously higher quality (see Sect. 4.2). Star identification based on the quality factor score estimated in Model 1 yielded a percentage of stars of approximately 1%. However, when combining this quality score with quantity, the percentage of stars was at a minimum of 0.15%. This result was expected because quality as measured in Model 1 by the latent variable η is disentangled from quantity (i.e., quantity and quality are uncorrelated).

We checked the Jaccard similarity of the pools of identified stars for all pairs of applied approaches (Table 8). Using the latent variable η for identification reports significantly smaller similarity scores with respect to the other criteria, indicating that this approach is based on different information. For example, the identification based on quantity and quality in terms forward citations is more similar to the group of stars based on quantity alone than the combination from quantity and quality defined by η . These similarities revealed that common practice to identify stars based on approaches that ignore the inherent relation of quantity and quality as implied by the EOB were indeed more alike as compared to the similarities of these approaches with star identification based on estimates of η . In other words, being identified as a star becomes less likely when the EOB is explicitly considered in a theoretically driven measurement approach.

Conclusion

In this work we extended the EOB, which accounts for the intricate relationship between quantity and quality in scientific productivity, into a multidimensional model that provides a practical assessment framework for the identification of star inventors.

Previous findings demonstrated that the EOB fits reasonably well to data on scientific productivity. However, these results were mostly limited to citation counts as quality indicators, whereas in this work quality of patents was operationalized as a multifaceted construct (i.e., we used citations, geographical scope, technological scope, generality, and originality as indicators). The multifaceted EOB proposed in this work (i.e., without modeling quality as a latent variable) fitted quite well to the data which provides further empirical support of the generalizability of the EOB to other quality indicators than citation counts. Furthermore, the analyses revealed that the residual variance conceptualized within the EOB framework was clearly larger than what is expected because of mere chance fluctuations. This finding was robust across the studied quality indicators and represents a prerequisite to model individual differences in hit-ratios as a function of latent variables.

In relation to this, it should be noted that besides the EOB other chance models (i.e., models that propose a random occurrence of high-quality products throughout a career) of creative productivity exist such as the Q model (Janosov et al., 2020; Sinatra et al., 2016). The Q model has been shown to fit data of scientists (Janosov et al., 2020; Sinatra et al., 2016), as well as data of people working in the movie business, the music business, or book authors (Janosov et al., 2020). This clearly hints at the wide applicability of chance models beyond science and patent inventors. For all these other fields, however, multidimensional extensions as proposed and empirically examined in the current work have not yet been considered. The Q model models the quality $S_{i,a}$ of product a produced by person i as a multiplicative function of Q_i (i.e., the creator's capacity generate high-quality products) and p_a (i.e., the "luck" parameter or potential quality of a product). This highlights that the Q model is formulated at a finer level of aggregation (i.e., the level of products) as compared to the EOB (i.e., quality indicators are aggregated for each person across products).

Hence, the Q model cannot be as simply integrated into a SEM framework as compared to the EOB. Nonetheless, we argue that extending the Q model into a multidimensional framework would be a useful extension to be tested in future research too. For example, analogous to the question (studied in the current work) in how far quality indicators load on the same person latent variable, it would be quite interesting to know if the luck component of a product is shared across different quality indicators.

In accordance with the previous literature that considered patent value and technological complexity as different subdimensions of patent quality, we found that the model including the two dimensions as latent variables displayed better fit than the unidimensional model. Moreover, regressing the latent quality variables on gender and career length revealed no significant result for gender, but a small positive relationship between career length and quality in the unidimensional model, and technological complexity in the two-dimensional model respectively.

The usefulness and applicability of the proposed assessment framework is highlighted by the good reliability and FDI findings that imply that factor scores (i.e., estimates for inventors' capacity for high-quality patents) can be used in subsequent analyses and in practical assessment activities. SEM models can be directly used for extending the research on scientific productivity and the relationship among value, technological complexity and other variables. Within SEM, such estimated relationships (e.g., latent variable regression coefficients) account for unreliability of observed measures. Although further refinements of the two-dimensional model seem not absolutely necessary for research contexts, for practical assessment contexts the factor scores for the value latent variable were not reliable enough, having an FDI < 0.90 (Ferrando & Lorenzo-Seva, 2018). Hence, further indicators of patent value would improve the reliability of the measurement. This is particularly the case when stakeholders and/or evaluators put a strong weight on the measurement of value in practical high-stakes decisions.

Despite its theoretical soundness and comparably better model fit, the two-dimensional multifaceted EOB model seems to be not suited for practical assessment purposes without further improvements. For this reason, we illustrated the identification of star inventors based on the unidimensional multifaceted EOB model. Even if this model fit to the data is worse than the two-dimensional model, it displayed still adequate fit. The factor loadings in the unidimensional model weighted all indicators of technological complexity as stronger than the value indicators. Hence, the capacity to invent high-quality products as measured in this model is associated to technological complexity more than to value indicators. This finding seems to explain why the identification of star inventors according to the unidimensional multifaceted EOB appeared to be different from the other common approaches employed in the literature that ignore the EOB. Indeed, the overlap between the stars defined by patent quantity and the stars based on the count of highly cited inventions is larger than the overlap between those same prolific inventors and the stars identified via the unidimensional EOB approach.

Finally, we equip researchers and evaluators with an R script to replicate all the findings reported in this work. This might pave the way for other scholars to employ the multifaceted EOB in their research and practical assessment projects, once the fit of the EOB to the data and the presence of individual differences in the model residuals are evaluated as requisite for application of the multifaceted EOB.

Our work is not exempt from limitations. In particular, the model specification does not account for potential heterogeneity in the impact of the resources of patent assignees on quantity or quality. The assumption here is that working for a company with many or few resources provides a proportional impact both on quantity and quality, while the effect

on one of the two dimensions might be disproportionately larger. In particular, this is of relevance for two aspects: when an inventor changes employer and the new one is very different from the prior in terms of available resources; the geographical scope can be particularly affected by employers' resources (i.e., more than the other quality metrics). Future research could improve the analyses by examining when the role of employers is more impactful on the investigated dimensions of quantity and quality. Finally, the results suggest that the characterization of the value dimension could benefit from the introduction of additional variables which might be tested on specific subsamples, such as the number of renewals of granted patents, licensing or sales values for transacted patents, or the direct assessment of the economic relevance through a survey of inventors.

Appendix

This section reports the robustness checks of the model fit of the unidimensional and two-dimensional multifaceted EOB models (Fig. 3). The models were tested on 35 different subsamples, corresponding to the technological fields in the WIPO concordance table which associates IPC codes to technical areas. On the left side in Fig. 3, the robustness check for the unidimensional multifaceted EOB is depicted. With the TLI as an exception, it is clearly visible that all other fit indices were at least acceptable for most of the WIPOs. Reliability and FDI values were good to excellent for latent quality factor scores across all WIPOs. This picture of results was found to be slightly reversed for the two-dimensional multifaceted EOB (see right side in Fig. 3). Model fit was clearly generally better for this model, but reliability for the value factor was below the recommended cut-offs for almost all WIPOs. In addition, although not visible in Fig. 3, technical estimation issues (e.g., Heywood-cases) occurred only for the two-dimensional model.

For detailed findings on each of the 35 subsamples, the interested reader can look at the Open Science Framework repository for this work (<https://osf.io/bjad4/>).

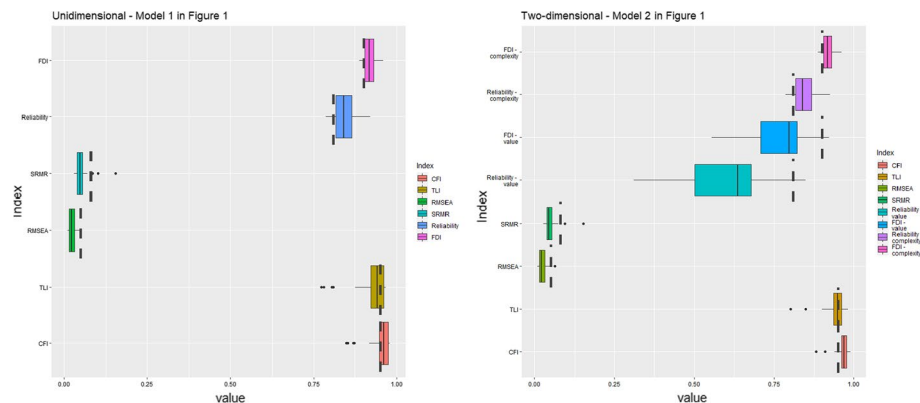


Fig. 3 Model fit, reliability, and FDI results summarized across all 35 WIPO technical fields for both the unidimensional Model 1 (left) and the two-dimensional Model 2 (right). Desirable cut-offs are depicted as dark gray dashed vertical lines. FDI, reliability, TLI, and CFI should be right to the cut-off, whereas SRMR and RMSEA should be left to the cut-off

Acknowledgements This work is a substantially extended version of our conference paper presented at the 18th International Conference on Scientometrics and Informetrics—ISSI 2021 (Cavaggioli and Forthmann 2021).

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Agostini, L., Cavaggioli, F., Filippini, R., & Nosella, A. (2015). Does patenting influence SME sales performance? A quantity and quality analysis of patents in Northern Italy. *European Journal of Innovation Management*. <https://doi.org/10.1108/EJIM-07-2013-0071>
- Agrawal, A., McHale, J., & Oettl, A. (2017). How stars matter: Recruiting and peer effects in evolutionary biology. *Research Policy*, 46(4), 853–867. <https://doi.org/10.1016/j.respol.2017.02.007>
- Aguinis, H., & O'Boyle, E. (2014). Star performers in twenty-first century organizations. *Personnel Psychology*, 67(2), 313–350. <https://doi.org/10.1111/peps.12054>
- Azoulay, P., Zivin, J. S. G., & Wang, J. (2010). Superstar extinction. *Quarterly Journal of Economics*, 125(2), 549–589. <https://doi.org/10.1162/qjec.2010.125.2.549>
- Beaudry, C., & Schifflauerova, A. (2011). Impacts of collaboration and network indicators on patent quality: The case of Canadian nanotechnology innovation. *European Management Journal*, 29(5), 362–376. <https://doi.org/10.1016/j.emj.2011.03.001>
- Bendersky, C., & Hays, N. A. (2012). Status conflict in groups. *Organization Science*, 23(2), 323–340. <https://doi.org/10.1287/orsc.1110.0734>
- Bergé, L., Carayol, N., & Roux, P. (2018). How do inventor networks affect urban invention? *Regional Science and Urban Economics*, 71, 137–162. <https://doi.org/10.1016/j.regsciurbeco.2018.05.002>
- Bollen, K. A. (1989). *Structural equations with latent variables*. John Wiley & Sons, Inc. <https://doi.org/10.1002/9781118619179>
- Brown, A., & Croudace, T. J. (2015). Scoring and estimating score precision using multidimensional IRT models. In S. P. Reise & D. A. Revicki (Eds.), *Multivariate applications series. Handbook of item response theory modeling: Applications to typical performance assessment*. Taylor & Francis Group.
- Call, M. L., Nyberg, A. J., & Thatcher, S. M. B. (2015). Stargazing: An integrative conceptual review, theoretical reconciliation, and extension for star employee research. *Journal of Applied Psychology*, 100(3), 623–640. <https://doi.org/10.1037/a0039100>
- Cavaggioli, F., & Forthmann, B. (2021). Star inventors: Quantity and quality in the EOB model. In W. Glänzel, S. Heffer, P.-S. Chi, & R. Rousseau (Eds.), *18th International Conference on Scientometrics and Informetrics, ISSI 2021* (pp. 217–228). International Society for Scientometrics and Informetrics. https://www.issi-society.org/proceedings/issi_2021/Proceedings_ISSI_2021.pdf.
- Cavaggioli, F., Colombelli, A., De Marco, A., & Paolucci, E. (2020a). How venture capitalists evaluate young innovative company patent portfolios: Empirical evidence from Europe. *International Journal of Entrepreneurial Behaviour and Research*, 26(4), 695–721. <https://doi.org/10.1108/IJEBR-10-2018-0692>
- Cavaggioli, F., De Marco, A., Montobbio, F., & Ughetto, E. (2020b). The licensing and selling of inventions by US universities. *Technological Forecasting and Social Change*. <https://doi.org/10.1016/j.techfore.2020.120189>
- Cavaggioli, F., & Ughetto, E. (2016). Buyers in the patent auction market: Opening the black box of patent acquisitions by non-practicing entities. *Technological Forecasting and Social Change*, 104, 122–132. <https://doi.org/10.1016/j.techfore.2015.11.031>
- Chen, J. S., & Garg, P. (2018). Dancing with the stars: Benefits of a star employee's temporary absence for organizational performance. *Strategic Management Journal*, 39(5), 1239–1267. <https://doi.org/10.1002/smj.2758>

- Den Hartigh, R. J. R., Van Dijk, M. W. G., Steenbeek, H. W., & Van Geert, P. L. C. (2016). A dynamic network model to explain the development of excellent human performance. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2016.00532>
- Dennis, W. (1958). The age decrement in outstanding scientific contributions: Fact or artifact? *American Psychologist*, 13(8), 457–460. <https://doi.org/10.1037/h0048673>
- Estabrook, R., & Neale, M. (2013). A comparison of factor score estimation methods in the presence of missing data: Reliability and an application to nicotine dependence. *Multivariate Behavioral Research*, 48(1), 1–27. <https://doi.org/10.1080/00273171.2012.730072>
- Ferrando, P. J., & Lorenzo-Seva, U. (2018). Assessing the quality and appropriateness of factor solutions and factor score estimates in exploratory item factor analysis. *Educational and Psychological Measurement*, 78(5), 762–780. <https://doi.org/10.1177/0013164417719308>
- Fischer, J., Ritchie, E. G., & Hanspach, J. (2012). Academia's obsession with quantity. *Trends in Ecology & Evolution*, 27(9), 473–474. <https://doi.org/10.1016/j.tree.2012.05.010>
- Forthmann, B., Leveling, M., Dong, Y., & Dumas, D. (2020). Investigating the quantity–quality relationship in scientific creativity: An empirical examination of expected residual variance and the tilted funnel hypothesis. *Scientometrics*. <https://doi.org/10.1007/s11192-020-03571-w>
- Forthmann, B., Szardenings, C., & Dumas, D. (2021a). On the conceptual overlap between the fluency contamination effect in divergent thinking scores and the chance view on scientific creativity. *The Journal of Creative Behavior*, 55(1), 268–275. <https://doi.org/10.1002/jobc.445>
- Forthmann, B., Szardenings, C., & Dumas, D. (2021b). Testing equal odds in creativity research. *Psychology of Aesthetics, Creativity, and the Arts*, 15(2), 324–339. <https://doi.org/10.1037/aca0000294>
- Forthmann, B., Szardenings, C., Dumas, D., & Feist, G. J. (2021). Strict equal odds: A useful reference to study the relationship between quality and quantity. *Creativity Research Journal*. <https://doi.org/10.1080/10400419.2020.1827605>
- Green, B. F., Bock, R. D., Humphreys, L. G., Linn, R. L., & Reckase, M. D. (1984). Technical guidelines for assessing computerized adaptive tests. *Journal of Educational Measurement*, 21(4), 347–360. <https://doi.org/10.1111/j.1745-3984.1984.tb01039.x>
- Groysberg, B., & Lee, L. E. (2009). Hiring stars and their colleagues: Exploration and exploitation in professional service firms. *Organization Science*, 20(4), 740–758. <https://doi.org/10.1287/orsc.1090.0430>
- Groysberg, B., Polzer, J. T., & Elfenbein, H. A. (2011). Too many cooks spoil the broth: How high-status individuals decrease group effectiveness. *Organization Science*, 22(3), 722–737. <https://doi.org/10.1287/orsc.1100.0547>
- Guan, J. C., & Gao, X. (2009). Exploring the h-index at patent level. *Journal of the American Society for Information Science and Technology*, 60(1), 35–40. <https://doi.org/10.1002/asi.20954>
- Guilford, J. P. (1968). *Creativity, intelligence and their educational implications*. Null (Vol. null).
- Han, S., & Ravid, S. A. (2020). Star turnover and the value of human capital-evidence from Broadway shows. *Management Science*, 66(2), 958–978. <https://doi.org/10.1287/mnsc.2018.3177>
- Harhoff, D., Scherer, F. M., & Vopel, K. (2003). Citations, family size, opposition, and the value of patent rights. *Research Policy*, 32, 1343–1363.
- Hess, A. M., & Rothaermel, F. T. (2011). When are assets complementary? Star scientists, strategic alliances, and innovation in the pharmaceutical industry. *Strategic Management Journal*, 32(8), 895–909. <https://doi.org/10.1002/smj.916>
- Higham, K., de Rassenfosse, G., & Jaffe, A. (2020). Patent quality: Towards a systematic framework for analysis and measurement. *Cambridge, MA*. <https://doi.org/10.3386/w27598>
- Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings of the National Academy of Sciences*, 102(46), 16569–16572. <https://doi.org/10.1073/pnas.0507655102>
- Hohberger, J. (2016). Does it pay to stand on the shoulders of giants? An analysis of the inventions of star inventors in the biotechnology sector. *Research Policy*, 45(3), 682–698. <https://doi.org/10.1016/j.respol.2015.12.003>
- Janosov, M., Battiston, F., & Sinatra, R. (2020). Success and luck in creative careers. *EPJ Data Science*, 9(1), 9. <https://doi.org/10.1140/epjds/s13688-020-00227-w>
- Kehoe, R. R., Lepak, D. P., & Bentley, F. S. (2018). Let's call a star a star: Task performance, external status, and exceptional contributors in organizations. *Journal of Management*, 44(5), 1848–1872. <https://doi.org/10.1177/0149206316628644>
- Kehoe, R. R., & Tzabbar, D. (2015). Lighting the way or stealing the shine? An examination of the duality in star scientists' effects on firm innovative performance. *Strategic Management Journal*, 36(5), 709–727. <https://doi.org/10.1002/smj.2240>
- Lacetera, N., Cockburn, I. M., & Henderson, R. (2004). Do firms change capabilities by hiring new people? A study of the adoption of science-based drug discovery. *Advances in Strategic Management*, 21, 133–159. [https://doi.org/10.1016/S0742-3322\(04\)21005-1](https://doi.org/10.1016/S0742-3322(04)21005-1)

- Lahiri, A., Pahnke, E. C., Howard, M. D., & Boeker, W. (2019). Collaboration and informal hierarchy in innovation teams: Product introductions in entrepreneurial ventures. *Strategic Entrepreneurship Journal*, 13(3), 326–358. <https://doi.org/10.1002/sej.1331>
- Lanjouw, J. O., Pakes, A., & Putnam, J. (1998). How to count patents and value intellectual property: The uses of patent renewal and application data. *Journal of Industrial Economics*, 46(4), 405–432. <https://doi.org/10.1111/1467-6451.00081>
- Lanjouw, J. O., & Schankerman, M. (2004). Patent quality and research productivity: Measuring innovation with multiple indicators. *The Economic Journal*, 114, 441–465.
- Lerner, J. (1994). The importance of patent scope: An empirical analysis. *The RAND Journal of Economics*, 25(2), 319. <https://doi.org/10.2307/2555833>
- Liu, K. (2014). Human capital, social collaboration, and patent renewal within U.S. pharmaceutical firms. *Journal of Management*, 40(2), 616–636. <https://doi.org/10.1177/0149206313511117>
- Lockwood, P., & Kunda, Z. (1997). Superstars and me: Predicting the impact of role models on the self. *Journal of Personality and Social Psychology*, 73(1), 91–103. <https://doi.org/10.1037/0022-3514.73.1.91>
- Martínez, C. (2011). Patent families: When do different definitions really matter? *Scientometrics*, 86(1), 39–63. <https://doi.org/10.1007/s11192-010-0251-3>
- Michalska-Smith, M. J., & Allesina, S. (2017). And, not or: Quality, quantity in scientific publishing. *PLoS ONE*, 12(6), e0178074. <https://doi.org/10.1371/journal.pone.0178074>
- Motiwala, M., Kumar, R., Ajmera, S., Lillard, J., Saad, H., Schultz, A., et al. (2020). Innovation, royalties, and introduction of the patent Hirsch index within U.S. Academic neurosurgery. *World Neurosurgery*, 137, e395–e405. <https://doi.org/10.1016/j.wneu.2020.01.223>
- Nijstad, B. A., De Dreu, C. K. W., Rietzschel, E. F., & Baas, M. (2010). The dual pathway to creativity model: Creative ideation as a function of flexibility and persistence. *European Review of Social Psychology*, 21(1), 34–77. <https://doi.org/10.1080/10463281003765323>
- OECD Patent Statistics Manual. (2009). *OECD Patent Statistics Manual*. <https://doi.org/10.1787/9789264056442-en>
- Office of the Chief Economist. (2019). Progress and Potential: A profile of women inventors on U.S. patents. *IP Data Highlights*, 2, 84–86.
- Oldroyd, J. B., & Morris, S. S. (2012). Catching falling stars: A human resource response to social capital's detrimental effect of information overload on star employees. *Academy of Management Review*, 37(3), 396–418. <https://doi.org/10.5465/amr.2010.0403>
- Prathap, G. (2018). Eugene Garfield: From the metrics of science to the science of metrics. *Scientometrics*, 114(2), 637–650. <https://doi.org/10.1007/s11192-017-2525-5>
- Rosseel, Y. (2012). lavaan : An R package for structural equation modeling. *Journal of Statistical Software*. <https://doi.org/10.18637/jss.v048.i02>
- Rothaermel, F. T., & Hess, A. M. (2007). Building dynamic capabilities: Innovation driven by individual-, firm-, and network-level effects. *Organization Science*, 18(6), 898–921. <https://doi.org/10.1287/orsc.1070.0291>
- Rousseau, R. (2018). The repeat rate: From Hirschman to Stirling. *Scientometrics*, 116(1), 645–653. <https://doi.org/10.1007/s11192-018-2724-8>
- Schmoch, U. (2008). *Concept of a technology classification for Country comparison. Final report to the World Intellectual Property Organization*.
- Simonton, D. K. (1988). *Scientific genius: A psychology of science*. Cambridge University Press.
- Simonton, D. K. (1988). *Scientific genius: a psychology of science*. Cambridge University Press.
- Simonton, Dean Keith. (1988). *Scientific genius: A psychology of science*. Cambridge University Press.
- Simonton, D. K. (2004). *Creativity in science: Chance, logic, genius, and zeitgeist*. Cambridge University Press.
- Simonton, D. K. (2009). Scientific creativity as a combinatorial process: The chance baseline. In P. Meusburger, J. Funke, & E. Wunder (Eds.), *Milieus of creativity* (Vol. 2, pp. 39–51). Dordrecht: Springer.
- Simonton, D. K. (2003). Scientific creativity as constrained stochastic behavior: The integration of product, person, and process perspectives. *Psychological Bulletin*, 129(4), 475–494. <https://doi.org/10.1037/0033-2909.129.4.475>
- Simonton, D. K. (2010). Creative thought as blind-variation and selective-retention: Combinatorial models of exceptional creativity. *Physics of Life Reviews*, 7(2), 156–179. <https://doi.org/10.1016/j.phrv.2010.02.002>
- Sinatra, R., Wang, D., Deville, P., Song, C., & Barabasi, A.-L. (2016). Quantifying the evolution of individual scientific impact. *Science*, 354(6312), aaf5239–aaf5239. <https://doi.org/10.1126/science.aaf5239>

- Singh, K. (1995). The impact of technological complexity and interfirm cooperation on business survival. *Academy of Management Proceedings*, 1995(1), 67–71. <https://doi.org/10.5465/ambpp.1995.17536285>
- Subramanian, A. M., Lim, K., & Soh, P. H. (2013). When birds of a feather don't flock together: Different scientists and the roles they play in biotech R&D alliances. *Research Policy*, 42(3), 595–612. <https://doi.org/10.1016/j.respol.2012.12.002>
- Swaab, R. I., Schaerer, M., Anicich, E. M., Ronay, R., & Galinsky, A. D. (2014). The too-much-talent effect: Team interdependence determines when more talent is too much or not enough. *Psychological Science*, 25(8), 1581–1591. <https://doi.org/10.1177/0956797614537280>
- Torrisi, S., Gambardella, A., Giuri, P., Harhoff, D., Hoisl, K., & Mariani, M. (2016). Used, blocking and sleeping patents: Empirical evidence from a large-scale inventor survey. *Research Policy*, 45(7), 1374–1385. <https://doi.org/10.1016/j.respol.2016.03.021>
- Trajtenberg, M., Henderson, R., & Jaffe, A. (1997). University versus corporate patents: A window on the basicness of invention. *Economics of Innovation and New Technology*, 5(1), 19–50. <https://doi.org/10.1080/10438599700000006>
- van Zeebroeck, N. (2011). The puzzle of patent value indicators. *Economics of Innovation and New Technology*, 20(1), 33–62. <https://doi.org/10.1080/10438590903038256>
- van Zeebroeck, N., & van de la Pottelsberghe Potterie, B. (2011). The vulnerability of patent value determinants. *Economics of Innovation and New Technology*, 20(3), 283–308. <https://doi.org/10.1080/10438591003668638>
- Wang, Y., Zhou, Z., & Li-Ying, J. (2013). The impact of licensed-knowledge attributes on the innovation performance of licensee firms: Evidence from the Chinese electronic industry. *The Journal of Technology Transfer*, 38(5), 699–715. <https://doi.org/10.1007/s10961-012-9260-0>
- West, S. G., Taylor, A. B., & Wu, W. (2012). Model fit and model selection in structural equation modeling. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 209–231). The Guilford Press.
- Zucker, L. G., & Darby, M. R. (1997). Individual action and the demand for institutions: Star scientists and institutional transformation. *American Behavioral Scientist*, 40(4), 502–513. <https://doi.org/10.1177/0002764297040004012>
- Zucker, L. G., Darby, M. R., & Armstrong, J. S. (2002). Commercializing knowledge: University science, knowledge capture, and firm performance in biotechnology. *Management Science*, 48(1), 138–153. <https://doi.org/10.1287/mnsc.48.1.138.14274>