Article

What Is Quality in Research? Building a Framework of Design, Process and Impact Attributes and Evaluation Perspectives

Alessandro Margherita *[ID]**, Gianluca Elia** [ID] **and Claudio Petti** [ID]

Department of Engineering for Innovation, University of Salento, 73100 Lecce, Italy;
gianluca.elia@unisalento.it (G.E.); claudio.petti@unisalento.it (C.P.)
* Correspondence: alessandro.margherita@unisalento.it; Tel.: +39-0832-297922

Abstract: The strategic relevance of innovation and scientific research has amplified the attention towards the definition of quality in research practice. However, despite the proliferation of evaluation metrics and procedures, there is a need to go beyond bibliometric approaches and to identify, more explicitly, what constitutes good research and which are its driving factors or determinants. This article reviews specialized research policy, science policy and scientometrics literature to extract critical dimensions associated with research quality as presented in a vast although fragmented theory background. A literature-derived framework of research quality attributes is, thus, obtained, which is subject to an expert feedback process, involving scholars and practitioners in the fields of research policy and evaluation. The results are represented by a structured taxonomy of 66 quality attributes providing a systemic definition of research quality. The attributes are aggregated into a three-dimensional framework encompassing research design (ex ante), research process (in-process) and research impact (ex post) perspectives. The main value of the study is to propose a literature-derived and comprehensive inventory of quality attributes and perspectives of evaluation. The findings can support further theoretical developments and research policy discussions on the ultimate drivers of quality and impact of scientific research. The framework can be also useful to design new exercises or procedures of research evaluation based on a multidimensional view of quality.

Keywords: design; evaluation; impact; quality attributes; research process; scientific research

1. Introduction

One of this paper's authors recently attended a conference at a university that displayed the slogan "University of fundamental excellence" on its walls and in its science park dedicated to impact and innovation. The slogan clearly meant to signal ambitions concerning the quality of the institution, and 'excellence' on its own seemed to be insufficient. This university is by no means unique, and we use the observation as an indication of the prevailing problems in framing the issue of the quality of scientific work, particularly in a time where research increasingly is expected to confine equally to ideas of "excellence" and "impact". Of course, excellence is driven by a number of multiple aspects and dimensions related to the quality of the education system, the implementation of sustainable development practices, the actual impact of the third mission, the ability to attract and welcome foreign students and many other factors.

In this article, we address the timely, extremely debated topic of what represents research quality and what characterizes quality research. We focus on building a multidimensional understanding of quality in research practice and an actionable framework of generally recognized concepts associated with the quality of scientific research. The immense societal attention to the strategic relevance of scientific research has amplified national and international systems of research evaluation [1]. Partly driven by global trends, national governments' perspectives on research quality have converged around rather vague notions of research excellence, which underpin many evaluation regimes [2]. In a

time when quality evaluation mostly happens implicitly through peer review or explicitly through the development of primarily bibliometric indicators, a systematic review of the many dimensions of quality is crucial. This paper aims primarily to contribute to our understanding of the multiple dimensions of quality in scientific research.

A number of factors contribute to the importance of evaluating the quality of scientific research today. First, the limits to financial resources available to universities and research institutions require a more efficient allocation of funds to research projects, groups and initiatives. Second, the international competitive context has caused, particularly in advanced countries, an evaluation fever and the development of global rankings of research entities [3,4]. Third, increased social awareness and attention of the public towards the process and the results of scientific research—not least in the light of grand challenges and with the coronavirus particularly in mind—stresses the need to document the activity of scientists and researchers and the overall impact of their actions for individuals and societies. Fourth, the quality of education and research is crucial for attracting international talents and innovative organizations willing to invest to create the conditions for territorial development.

The evaluation of research processes and outputs, such as publications, grants and promotion decisions, research projects, spin-offs, patents and initiatives able to address social issues and development of societies, is mostly performed by peer review, yet increasingly supported by bibliometric information derived from basic indicators such as number of publications, number of citations and the journal's impact factor. Indicators are used as a "proxy" measure of research quality and productivity and are largely adopted as a basis for orienting academic and policy decisions (e.g., career advancements or budget allocation). There is a long-standing debate about the limitations and biases of the quantitative measures of quality, including disciplinary differences, social and circular effects and the shortcomings of output indicators [5–9]. The current state-of-the-art in research evaluation is observed to consist of a combination of quantitative and qualitative measures [10], but there is a need to explore what a fruitful combination may entail by looking more closely at the dimensions of quality.

The importance of moving beyond bibliometrics and of developing more holistic approaches to evaluating research quality has triggered a worldwide debate on the ways in which the output of scientific research is evaluated by funding agencies, academic institutions and other parties. In 2012, a group of editors and publishers of scholarly journals created the San Francisco Declaration on Research Assessment (DORA), which provided a number of recommendations, such as to consider a broad range of measures including qualitative indicators of research impact (e.g., influence on policy and practice). In 2014, the European Commission started an online public Consultation on "Science 2.0": Science in Transition about the changing science system. It has highlighted the limitations of traditional metrics and the need to develop alternative methods to monitor (open) science activities. The Leiden Manifesto for Research Metrics from 2015 provided principles for the measurement of research performance and more sustainable and comprehensive approaches to research evaluation [11]. Finally, the UK review on quantitative indicators, *The Metric Tide*, explored the use of metrics across different disciplines and assessed their potential contribution and limitations to the development of research excellence and impact. It emphasized that metrics should be "responsible", i.e., robust, transparent, dynamic and based on a recognition that quantitative evaluation should support qualitative and expert assessment rather than the other way around [12].

Although these initiatives have highlighted that quantitative evaluations should be increasingly integrated with qualitative evaluations, it is extremely complex to define rigorous methodologies and tools that can be applied to capture the broader value of scientific research. There is tension between defining simple (but invalid) indicators that are widely used and more sophisticated indicators that cannot be used because they are not transparent, cannot be calculated or are difficult to interpret [13]. Moreover, the use of different measures makes it difficult to compare an institution's evaluation results with those of

other institutions or disciplines. The available scientific literature on research quality, and on what can be defined as research, is scarce, and there is, thus, a need for determining a systemic definition of the quality of research practice [14]. In particular, there is room for new contributions attempting to define and discuss the multidimensional meaning of quality into scientific research and the strategies through which it can be measured. Additionally, it is of interest to identify more practical and fine-grained evaluation criteria and metrics to be applied into official or informal assessment exercises.

It is difficult to find a general definition of what constitutes good scientific practice and research quality in the international community. Initiatives have been taken in university rankings to develop systems that use multiple dimensions and, thus, several potential rankings [4]. There is a need to develop similar approaches for evaluating the quality of scientific research at lower levels, for example, in the evaluation of funding proposals, manuscripts and people. The ultimate goal would be to build a dashboard or suite of indicators [15] able to address both conventional outcome measurements and evaluations of scientific leadership and citizenship. While many contributions in the literature have focused on specific definitions of quality in science and research, an integrative framework gathering all the perspectives is still missing.

There is, in particular, a need to move beyond the rhetoric of excellence, which has become the dominant perspective on quality in academic work and policy. Excellence is today's gold standard of the university world and a holy grail of academic life [16]. The concept of "excellence" is pervasive across the academy and is used to refer to research outputs as well as to researchers, theory and education, individuals and organizations. The authors of [17] have argued that the term has no intrinsic meaning but functions as a linguistic interchange mechanism. An unqualified emphasis on excellence could undermine the very foundations of good research and scholarship and the standards they build upon. In addition, what is meant by excellence differs a lot between stakeholders and disciplines in research.

Both for academic and policy purposes, a more nuanced perspective is needed on quality as "academic standards" rather than "excellence". In such a perspective, the "U-Multirank" initiative (https://www.umultirank.org, accessed on 9 February 2022) is a good example of an effort aimed to build an international ranking of comparable higher education institutions along different dimensions of activity and to allow users to develop personalized rankings by selecting indicators in terms of their own preferences.

The main research gap addressed by this article is, thus, the lack of a systemic definition of research quality dimensions, providing a more nuanced and cross-disciplinary understanding of what constitutes good research and its driving factors. At this purpose, we conduct a systematic literature review and gather expert feedback to derive and validate an integrative framework of design, process and impact attributes and evaluation perspectives. The remainder of the work is structured as follows. We start (Section 2) with a review of relevant literature and theoretical perspectives, leading to an account of our own research process (Section 3) undertaken to gather quality dimensions and attributes from the literature. Next, we present the main results of this research process. First, we illustrate (Section 4) the attributes of quality in scientific research. Next, we present (Section 5) an integrative process description. We then discuss (Section 6) the advancements with respect to extant theory as well as a number of policy implications, limitations and avenues for further research.

2. Background

Historically, quantitative evaluation of scientific research quality and productivity has been based in particular on counting and analyzing publications and their received citations [18,19] as well as related measures such as the journal's impact factor [20] and the author's h-index [21]. Based on this foundation, other indicators have been introduced to reduce bias and to measure the scientific "core" output of a researcher [22] such as the g-index [23] and the hg-index [24]. In the last ten years, the demand for bibliometric

assessment has resulted in a further adaption of new indicators or variants/combinations of established ones. The authors of [25] categorized 108 indicators (including publication count, output, effect of output and individual ranking indicators), whereas [26] provided an in-depth review of the literature on citation impact indicators. As such, there is no lack of quantitative indicators, and they have been theorized and used for at least half a century. Bibliometric indicators are judgment devices [27], which can render the evaluation process more efficient and cost-effective.

However, research assessment contexts also involve expert judgement in the form of peer review, and state-of-the-art evaluation approaches combine quantitative indicators and qualitative expert judgement [10]. The evaluation of specific research or of a researcher is evaluation of something "unique", which requires an approach that allows the valuation of "singularities" [28]. It has been argued that the future of research evaluation rests with an "intelligent combination" of advanced metrics and transparent peer review [29,30]. A relevant attempt to integrate "traditional" quantitative and bibliometrics-related metrics with more qualitative and individual-specific indicators was provided by Holbrook and colleagues, resulting in a list of 56 indicators that combines quality and external impact and also seeks to highlight "negative" impacts and phenomena [31]. Examples include impact indicators associated with public engagement (e.g., participation in public education, mention by policy makers and public research discussions), indicators associated with the academic community (e.g., interdisciplinary achievements and faculty recommendations) and indicators associated with the media (e.g., social-network contacts or website hits). Another initiative is Snowball (https://www.snowballmetrics.com, accessed on 9 February 2022), a bottom-up project started by an international group of research-intensive universities as an alternative to metrics developed by governments and funding organizations. Snowball and the lists in [31] are examples of proposed indicator systems that are tailored to be combined with peer review and other forms of qualitative judgement.

The tendency to shape evaluation criteria and systems towards societal contribution has been accelerated by claims that knowledge production itself is changing where societal needs and perspectives increasingly become internalized in research. One example is "Mode 2" [32], arguing that traditional disciplinary research (Mode 1) is increasingly replaced by transdisciplinary knowledge production in the context of application (Mode 2), and the perspective has been used to study quality dimensions across multiple sectors and disciplines [33,34]. Leading evaluation scholars have argued that science policy needs to move beyond an emphasis on commercial outputs and embrace wider intellectual, social, cultural, environmental and economic returns where qualitative measures and processes can be highlighted [10]. However, there are also growing concerns about the quality and effectiveness of the peer review system, and some authors have proposed new procedures and technologies, e.g., to flag problematic publications [35].

Another relevant trend in the assessment of research quality is the increasing relevance of the web and social networks in showcasing, documenting and measuring the activities and results of scientists and researchers. Under headings such as "The future of bibliometrics", "Next-generation metrics" and "Web 2.0" [36–38], new web-based indicators such as mentions, acknowledgments, endorsements, downloads, recommendations, blog posts and tweets are discussed. Such alternative metrics, or "altmetrics", show a number of potential advantages such as openness, quick accumulation and real-time traceability of a large variety of research outputs. These web-created metrics (webometrics) allow incorporating the impact of the web on the influence exerted by researchers and scientists on the online and offline community (influmetrics). However, altmetrics also show a number of risks such as an effect of "commercialization"; several data quality biases; a lack of theory and empirical evidence; and the risk of gaming and manipulation [39].

The evaluation of societal impact is more complicated than evaluating academic quality or impact alone, i.e., an intellectual contribution [40]. One particular problem is to find adequate tools and methods to measure impact and disentangle the extent to which the research results are the sole (or most significant) causes of the effect produced, which is known

as the attribution problem in evaluation [10,40,41]. This is where qualitative judgement may be especially needed or useful. Definitions are also challenging, and there is a lack of consensus on the meaning of the words social and societal and the methods for measuring social, cultural, environmental and economic returns from publicly funded research [42,43]. With the ultimate goal to define social impact, different concepts were introduced such as "third-stream activities" [44], societal benefits [45], usefulness/utility [46,47], public values [48,49], knowledge transfer [4] and societal relevance [50,51]. The social impact of research occurs when disseminated results produce improvements in relation to the goals of society, such as socio-economic cohesion, quality of life, employment, human capital formation, public health and security. Evaluations often distinguish between outputs (texts, patents, objects, etc.), uptake (engagement with research activity by users), use (discussion, sharing and application of results) and impact (changes in awareness, knowledge, ideas, attitudes, policy and practice), but many of the traditional quality indicators focus on outputs alone [52].

The attention towards impact concerns has created a new trend into institutional policies as well as into (peer) review of grant proposals submitted to public science funding bodies [53]. Ref. [51] compared the procedures of the US National Science Foundation and the European Commission's 7th Framework Programme for evaluating, ex ante, the potential societal impact of research proposals. The UK Economic and Social Research Council highlighted the relevance of meeting impact expectations (besides innovation and interdisciplinarity) in the design of high-quality research initiatives. For the European Research Council [54], two major areas of evaluation of research projects are the use of interdisciplinary approaches as a strategy to solve complex problems and the potential to generate a sustainable socio-technical impact. In a study exploring the perceptions of evaluators of the UK Research Excellence Framework (REF) 2014, the criteria identified as relevant for assessing impact and quality of research outputs are originality, significance and rigor [55].

Quantitative evaluations of scientific quality are heavily rooted in bibliometric indicators and evaluation systems. These are powerful but have major limitations for coming to grips with the complicated notion of "quality", and they are, therefore, normally combined with peer review or expert/qualitative judgement. A number of examples of combined evaluation systems exist, but none of them has emerged as a new standard. A further problem is that the term quality now often includes impacts beyond research, which is a very complicated issue in itself. To us, this indicates that it is important to move beyond indicators or expressions of the importance of peer review and to maintain a fundamental debate about the underlying dimensions of research quality. This paper aims to contribute to the debate by systematically looking at how quality is elaborated in the literature and then by confronting a set of experts with these results.

3. Research Process and Method

The particular type of literature review process undertaken in this paper was separated into four steps, as represented in Figure 1. In Stage 1 (Review of Literature), we focused on reviewing cross-disciplinary specialized literature on research policy, science policy and scientometrics, with a specific attention on studies investigating the multi-sided meaning of quality in scientific research and attempts to describe and systematize the same. We used Google Scholar®, ISI WoK® and Scopus® databases to search the strings "research quality", "quality of scientific research", "quality of research", "research evaluation", "quality evaluation" and "quality assessment" in the titles, abstracts and keywords of research articles. After a first refinement, based on the reading of abstracts to exclude non-pertinent works, we selected 93 research articles for more in-depth analyses.
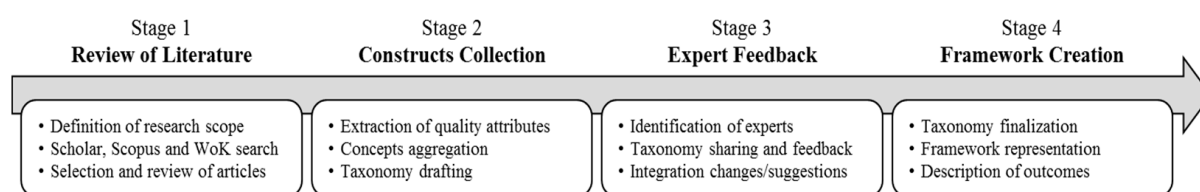
Figure 1. Research process undertaken to build the integrative quality framework.

In Stage 2 (Constructs Collection), we extracted, from selected papers, all constructs mentioned by authors as attributes or dimensions of research quality. We initially annotated (into a datasheet) all keywords, terms, concepts and variables mentioned in the analyzed articles within definitions, classifications and conceptual models of research quality. We did not rename concepts but rather left authors' wording and own definitions. We thus created a long list of concepts found across all articles and then placed the same concepts into an alphabetical order. This allowed us to immediately identify identical or easily comparable terms (e.g., clarity and clearness), which were unified into a single concept. We did not separate quality constructs or attributes related or associated to different "objects" (e.g., proposals, articles, projects, researchers, centers, etc.) as our analysis aims to define a comprehensive and more general definition of quality to be applied along multiple perspectives and units of analysis. We, thus, obtained a draft taxonomy of quality dimensions based on a preliminary consolidation (e.g., elimination of duplicates and redundancies) and aggregation of common contributions. Table 1 shows the outcomes of concept extraction work from the literature.

Table 1. Attributes or dimensions of research quality found in literature.

Attributes	Authors
Clarity, rigor, methodological soundness and craftsmanship	[56,57]
Coherence	[58]
Collaboration distance between citing/cited authors	[59]
Communicability (consumability, accessibility and searchability)	[14]
Communication and collaboration	[60]
Openness, universalism, disinterestedness and organized skepticism	[56,61,62]
Conforming (ethics, alignment with rules and sustainability)	[14]
Contextualization	[63–65]
Contributory (originality, relevance and generalizability)	[14]
Credibility (rigorousness, consistency, coherence and transparency)	[14]
F(ield)-type and S(pace)-type quality	[66]
Intellectual influence	[67]
Internal validity, reliability and rigor	[68,69]
Journal impact factor, citations and H-Index	Various
Methods, intellectual and political/social significance and originality	[56]
Novelty, methodological soundness and significance	[70]
Originality/novelty	[16,56,57,61,62,71,72]
Plausibility/reliability	[56,61,62,71]

Table 1. *Cont.*

Attributes	Authors
Scholarly exchange, connecting to other research, impact on research community and future research, innovation, originality, productivity, rigor, fostering cultural memories, recognition, reflection, criticism, continuity, continuation, openness, variety, self-management, independence, scholarship/erudition, connection between research and teaching	[73]
Scientific quality, defined scope, anticipated outcomes, operationalization, feasibility, process evaluation, documentation and dissemination	[74]
Scientific value and societal relevance value	[62]
Scientific, technological, clinical and socio-economic significance	[75]
Significance, approach, innovation, investigators and environment	
Societal quality, usefulness	[45,46]
Socio-economic impact, resource attraction and resource management	
Stringent argumentation, presentation of relevant evidence, clear language and structure, reflection of method and adherence to standards of scientific honesty	[73]
Stringency, intra- and extra-scientific effects and breadth	[76,77]
Technological and social merit	[78]
Transparency	[79,80]
Continuity, innovation, originality, rigor, reflection, criticism, scientific exchange, inspiration, connection to society, diversity, variety, topicality, openness, integration, autonomy, productivity, intrinsic motivation, scholarship and connection with teaching	[81]
Truth/probability, testability, coherence, simplicity/completeness, honesty, openness and impartiality/objectivity, originality and relevance/fruitfulness/value and verisimilitude	[56,61,62,82]
Value/usefulness	[71]

In the third Stage (Expert Feedback), we submitted our draft list of quality dimensions to a panel of experts in the fields of research policy, research evaluation and scientometrics. We used journal websites to obtain the names and e-mail contacts of 12 editors, associate editors and other members of the editorial boards of leading journals focused on the above-mentioned topics. We then prepared a one-page document (which is reported in the Appendix A of this paper) containing the list of quality dimensions and attributes, along with a tentative clustering of the same into general categories (e.g., impact-related attributes, process-related attributes) and an enclosed letter to explain the rationale and the goal of the study. We sent e-mails to experts, requiring feedback in terms of the following: (a) utility of the research and suggestions to refine the overall purpose, also through new literature or practitioner evidence; (b) more specific integrations and amendments to the proposed list of quality dimensions; (c) association of quality dimensions to different categories (research design, research process and research impact).

We obtained answers from 7 experts. In particular, we received complete feedback on the points (a), (b) and (c) from 5 experts, who found the research interesting and useful in both academic and policy discussion, as well as a relevant knowledge platform for stimulating further definition and codification efforts. From the other 2 experts, we only received more general comments. One of the experts stated that, in addition to defining a long list of criteria, an interesting problem is related to "how to interpret each criterion in relation to specific instances and how are trade-offs between different criteria handled in specific situations ... an important phenomenon is "cognitive contextualization" which means that criteria are given different weight in different situations and across

disciplines". The most critical expert stated that there could be "ambiguity in the meaning of quality-related concepts and in the reduction of the same to a more limited number of dimensions . . . the conceptual framework may be thus useful in specific context in which there is a clear objective".

In Stage 4 (Framework Creation), we finalized the taxonomy, also using a cluster analysis approach to improve validity and ensure that the chosen concepts together provide a robust evaluation checklist. Overlaps and commonalities among quality dimensions were addressed and the final framework was obtained. The process represented a deeper engagement with the topic than what a traditional literature review would allow, with two major outcomes: (a) a literature-derived and expert-validated comprehensive list of attributes associated with research quality; and (b) an integrative framework of macro-categories of evaluation and key items. These results are presented in the next section.

4. Attributes of Quality in Scientific Research

There are many relevant contexts for evaluating and debating the quality of research, and these include organizational contexts or settings (e.g., universities versus research institutes), disciplinary contexts (e.g., science versus humanities) and specific objects of interest for evaluation (e.g., research outputs versus authors). A fundamental distinction, regardless of other aspects, is between ex ante, in-process and ex post considerations. Evaluating research means evaluating aspects that are related to all activities ranging from ideation to the design of research, proceeding with execution and reporting and ending up with publication and diffusion. This means that attributes and dimensions of research quality can be related to the following:

1. *Research Design:* All that relates to the conceptualization of research, its aims and specific goals, the strategy or approach adopted by the researcher, the initial assumptions and ultimate focus of research. This can be termed an ex ante approach to quality.
2. *Research Process:* The execution of research activities, the research method and tools applied, the conduct of the researcher and the formalization and reporting of results. While not a very common focus in most evaluations, this interim approach is the closest to the everyday practice of the researchers.
3. *Research Impact:* The sharing of results, the influence on scholars and practitioners, the adoption or utilization of findings and the ultimate effects on the society. Ex post approaches to quality differ in whether they are interested mainly in the impact within the research system or outside of it.

Design is specifically concerned with the goal and the focus of the research, which can span across different disciplines and areas of human knowledge (e.g., research on entrepreneurship attempting to investigate also the psychology and neuroscience foundations of risk aversion). The aspect of interdisciplinarity shows that research quality is often about expressions of stakeholder preferences rather than expressions of a scale from "good" to "bad" onto which all research can be placed. This means that cross-disciplinary research is not necessarily "better" than research that scores low on this criterion, but the use of interdisciplinary approaches is seen by many as a strategy to solve complex problems and to generate sustainable socio-technical impacts [57]. As such, it also represents an assumption of the relationship between ex ante and ex post characteristics—cross-disciplinary research is (by some) believed to possess a greater propensity to result in certain desirable outcomes.

Process is concerned with the data, methodology and reporting of research. The robustness of approaches and the tools adopted are crucial for ensuring replicability, rigorousness, validity, reliability and consistency of "credible" research [14]. Research should also be a "catalyst" of needs expressed by stakeholders, who should be involved early in requirements identification [58], although the scope of inclusion of stakeholders will vary considerably based on the field. Good research is a cost-effective combination of tangible and intangible outcomes and the type of effort undertaken depends on the desired "effect" or result, ranging from no appreciable contribution to incremental scientific contribution, up to scientific breakthrough. The method applied (strategy, protocol and

techniques) and used to design, conduct and monitor research efforts can be "traditional" or conventional methods or innovative ones.

Finally, Impact can be viewed as the personal, academic and social influence of the research and its outputs. Currently, the relations between science and society, and the scientific awareness of the public, are greater than ever before and this generates new pressures to provide evidence of how and how much science generates social impacts [41]. Such impact is often measured by using informetrics data (e.g., article citations, journal rankings and impact factor), altmetrics and acknowledgment measures will be used to assess if and how the research stands the test-of-time. The evaluation of the impact of research is a quite debated theme. An interesting issue is whether it is possible to have a high score on impact and a low score on quality, i.e., the relationship between this aspect, area or timing of quality and the ones dealing with design and process. Tables 2–4 show the literature-derived and expert-validated lists of attributes of research quality associated to the Design (D), Process (P) or Impact (I) angle of analysis. The full taxonomy includes 66 attributes related to research design (13 attributes), research process (31) and research impact (22).

Table 2. Attributes associated to Research Design (D).

1	Authors' distance	2	Investigator's expertise
3	Craftsmanship	4	Interdisciplinarity
5	Credibility	6	Resource attraction
7	Disinterestedness	8	International scope
9	Objectivity/Impartiality	10	Stringent argumentation
11	Systems view	12	Honesty
13	Organized skepticism		

Table 3. Attributes associated to Research Process (P).

14	Clarity	15	Originality of method
16	Coherence	17	Cost effectiveness
18	Communicability	19	Conformance to ethics
20	Reliability	21	Reproducibility
22	Rigorousness	23	Evidence disclosure
24	Methodological soundness	25	Structure clarity
26	Stakeholder involvement	27	Documented sources
28	Accessibility and intelligibility	29	Thoroughness
30	Adherence to standards/rules	31	Internal validity
32	Unconventionality	33	Operationalization
34	Testability	35	Universalism
36	Completeness	37	Consistency
38	Replication feasibility	39	Transparency
40	Expert feedback	41	Plausibility
42	Scope tailoring and focus	43	Research process evaluation
44	Truth and veridicity		

While the framework can be useful to support theory and policy discussion aimed at elaborating a comprehensive and shared understanding of quality in research practice, the practical application of the model in evaluation exercises needs further design efforts. In particular, it is relevant to highlight that a selection of criteria is the desirable choice when a specific evaluation context or purpose (or exercise) has to be addressed. In such perspectives, the proposed framework provides a comprehensive definition of quality-related criteria, which may be of higher or lower relevance according to the intended evaluation. Moreover, there is a need to identify information sources or evidence to support a qualitative or quantitative analysis and the evaluation of the different quality attributes identified above. Table 5 shows an illustrative identification of evidence or information to support evaluation.

Table 4. Attributes associated to Research Impact (I).

45	Political and social significance	46	Scholarly exchange/diffusion
47	Practitioner impact	48	Scientific significance and value
49	Social impact/significance	50	Economic impact/significance
51	Societal relevance and value	52	Educational impact
53	Community impact	54	Intellectual significance
55	Technological significance	56	Theory impact
57	Usefulness	58	Trendsetting and future outline
59	Novelty	60	Sustainability of outcomes
61	Generalizability	62	Relevance
63	Contextualization	64	Searchability
65	Openness	66	Dissemination potential

Table 5. Focus of quality and illustrative information or evidence to support evaluation.

Focus of Quality	Illustrative Information or Evidence
Research Design (Motivation, Concentration, Approach)	keywords, research classification, purpose, research problems/questions, literature/practitioner gap, researcher profile, assumptions, research team, research strategy
Research Process (Data, Method, Reporting)	information sources, data processing methods and tools, partnerships, research context, reports, research funding, research project, case study protocol, full-time equivalent, products/services
Research Impact (Personal, Academic, Social)	citations, H-index, journal impact factor, consulting activities, social initiatives, career advancements, community acknowledgment, web presence, positions, start-ups, new products/services, patents, policy documents, industry reports, training outcomes

5. An Integrative Process Description

A point that deserves further reflection is the analysis of quality foci and related attributes into a process view of research activities. In the previous section, we described research design, research process and research impact namely as "ex ante", "in-process", and "ex post" perspectives of research quality. The three perspectives can be mapped on what constitutes a process of scientific research and this may also allow identifying and discussing relations among the same perspectives. Using a quite general view, scientific research is described as a six-phase endeavor that includes the following:

1. *Ideation*, i.e., the identification of a rationale or motivation (triggering factor) for conducting research and the high-level goals of the same;
2. *Preparation*, i.e., the definition of strategy, detailed objectives and a plan to conduct research activities;
3. *Execution*, i.e., the actual undertaking of research activities;
4. *Reporting*, i.e., the writing and formalization of a story about what we performed and what results we achieved;
5. *Publication*, i.e., the sharing the outcomes of our research within the academic/practitioner community;
6. *Diffusion*, i.e., the post-publication dissemination of results and the deriving effects.

While design attributes of research quality can be mostly associated to the Ideation step, process attributes are explicitly concerned with Execution and Reporting, and impact attributes are associated with Diffusion. Preparation is concerned with both design and process attributes, whereas Publication relates to process and impact attributes. In addition to identifying links among quality foci and phases of the research process, it may be of

interest to discuss links among quality foci over a process perspective. Three links can be identified in this regard.

An implementation link can be identified between design and process attributes of research. The quality of research is indeed related to how the research design is translated into proper action. Although original research concept and robust design are defined, the quality of research will be hindered by a poor execution of planned activities. Second, an externalization link exists between research process and impact. Quality is associated with the potential of research to produce effects to be shared within a relevant community; thus, quality depends on the ability of the researcher to bring a constant look at how research can contribute to improving current understanding and behaviors. Finally. An effectiveness link can be described between research design and impact. The quality of research is, of course, anchored to the actual achievement of the designed goals, and this may be dependent from an effective design of how research activity has to be conceptualized and implemented. Figure 2 presents a snapshot of research process, quality attributes and links.
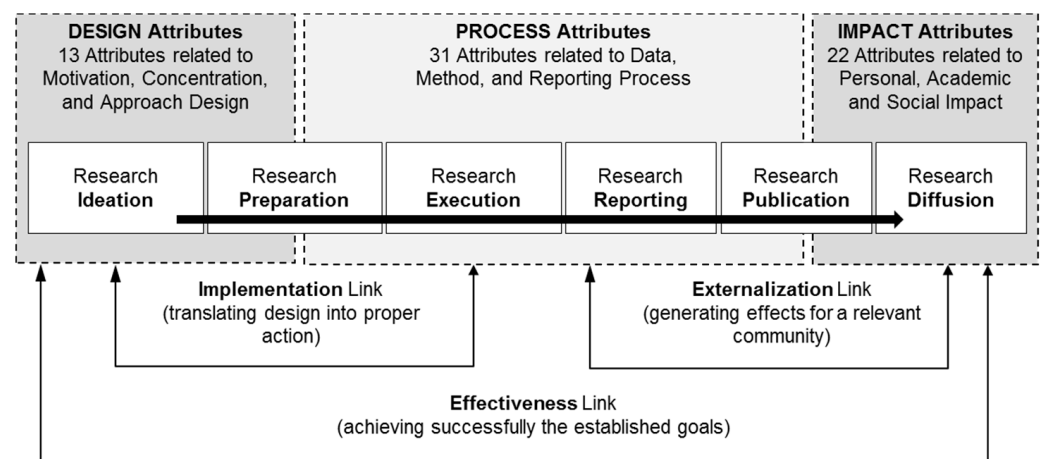

Figure 2. Integrated process description of research quality.

6. Discussion and Conclusions

6.1. Main Contribution and Comparison with Literature

We have presented a comprehensive framework of 66 attributes associated with the quality of scientific research, along with a classification of attributes based on a design, process and impact views of research quality. The framework is based on a systematic review of literature on research evaluation, research quality and impact, and it is refined using feedback of experts in the field of research policy and evaluation. For our best knowledge, the article represents the most extensive attempt to describe the meaning of quality in research in terms of a number of specific attributes. The framework provides a comprehensive approach to research evaluation as recommended in the emerging international debate (e.g., Leiden Manifesto, DORA).

Given the lack of an integrative definition in literature, the study has attempted to bring an inventory of the elements that constitute what we might refer to as "the essence" of research quality. The quality framework can serve as food for thought and a platform for further development of common concepts, terms and criteria associated with quality evaluations within and across specific research domains. More than attempting to suggest policy or normative definitions of what and how a component should be measured into scientific research, the framework may be used in an inclusive debate for the further development of relevant elements, weights and operationalization related to the quality of research practice in different academic fields.

Out study aims to systematize research quality dimensions in order to create broad frameworks for empirical research, policy development and evaluation. The study builds on recent contributions published by both research policy and interdisciplinary journals

(e.g., Minerva and Research Policy), aiming to define integrative and generalizable frameworks to describe what is quality in research and which are its driving factors or determinants. In particular, we were inspired by the purpose and the outcomes of the research conducted by [14]. Our study has some similarities with [14], who realized an extensive research effort to develop a model of research practice and to define the concepts related to its quality. Their paper is a platform study for the further development of criteria for research quality understanding and evaluation within and across research domains. Our research adds to [14] along two different perspectives, i.e., the process for building the model and the nature of findings, which also has implications for the utilization of the same framework in future studies. Concerning the model, [14] used working groups with researchers and modelling experts who contributed to building the proposed framework, while our paper is founded on an extensive review of specialized literature, which has allowed a more theory-derived model. We then used experts to refine and validate findings.

In terms of findings, [14] presented a hierarchy of 32 research quality concepts related to four areas labelled "Credible", "Contributory", "Communicable" and "Conforming". The four areas, thus, represent macro-characteristics of good research to which all 32 concepts can be associated. In our research, we present a more granular taxonomy of 66 attributes of research quality that have been classified using a stage-based and dynamic view of research. The attributes are associated with three groups of design (ex-ante), in-process and impact (ex-post) attributes or perspectives and such classification is potentially of easy operationalization as a dashboard to support evaluation exercises, which take the research process and its phases as the unit of analysis.

Our research can be also discussed in relation with the work of [66], who developed a novel framework to study and understand research quality across three key dimensions: first, by distinguishing between quality notions that originate in research fields (Field-type) and in research policy spaces (Space-type); second, by identifying in extant research three attributes considered important for good research, i.e., originality, reliability and value; and third, by defining five different "sites" where research quality concepts have relevance, i.e., researchers, knowledge communities, research organizations, funding agencies and national policy arenas. Our research adds to the extant contributions in that, although we share the logic of quality as a multifaceted notion, we provide an articulated inventory of quality attributes, thus specializing the broad areas of originality, reliability and value defined by the authors.

With respect to [31], who identified 56 metrics integrating "traditional" quantitative indicators and more qualitative and individual-specific indicators, we focused on attributes of research quality, which in turn can be used to define dashboards of purposeful and context-dependent metrics.

Research quality is a complex and multifaceted concept. Intellectual influence (for example measured by citations) and broader notions including "quality" and "excellence" do not necessarily coincide [83]. In our study, we presented a universal concept model that can have a general applicability and work as a menu from which more specifically focused evaluations can be composed, rather than a reference for all evaluation exercises covering the entire broad scope. The framework mostly has a summative priority for application in different situations and is intended to support further discussion on system approaches to quality evaluation. Evaluation is an important social activity for public organizations, and this is particularly true in the research field, where public resources are used to provide benefits for individuals and the society. Literature on evaluation systems and practices is extremely heterogeneous and comprises thousands of contributions from different journals and disciplines [84] and the worldwide debate is today on the importance of defining more integrative and responsible metrics for research evaluation [85]. The meaning of research quality varies with the context, but defining a comprehensive inventory of dimensions or attributes associated with quality may serve as an important starting point to build new evaluation frameworks.

In our research, we have analyzed specialized research policy, science policy and scientometrics literature to extract critical dimensions associated with research quality. The literature-derived framework was then submitted to expert scholars and practitioners in the fields of research policy and evaluation. The expert review allowed gathering suggestions in terms of integrations and amendments to the proposed framework. The interaction with experts allowed also obtaining insights and comments, which have been used in the definition of the main assumptions, the discussion of the potentialities and limitations of practical implementation and the overall limitations of our study.

6.2. Policy Implications and Practical Use of the Framework

The quality framework can support the analysis of research activity and results based on a set of crucial variables and parameters as well as the definition of a "research scorecard" or balanced scorecard [86] of research. The framework can provide a starting point and basis for discussion for the identification of key performance areas and the consequent planning of activities, goals to achieve and metrics to apply for measuring progress. The framework could also be used to support more quantitative assessment procedures. In this case, the framework can be used to address the set of assumptions, priorities and evaluation requirements that are defined for the specific evaluation context. For example, for some sectors or evaluation purposes, the degree of interdisciplinary exploration could be irrelevant whereas priority could be assigned to the research method adopted. The use of a quantitative approach could better serve the purposes of an "externally conducted" evaluation or assessment. However, the application of an evaluation algorithm to define and compare numerical results is a quite complex situation that deserves careful analysis to ensure applicability and reasonableness and to avoid issues of biases or evaluation distortions. The taxonomy of attributes reports widely shared characteristics of good research, with a corresponding set of criteria on the basis of which research could be evaluated, although not all conditions can be present in every research. A possible way to use the proposed framework could be generating a research profile (a graph) formed by judgments on the presence of different possible characteristics in the evaluated research. This can also be used to analyze whether funding or evaluation systems unfairly disadvantage certain types of research. Another potential use is to create a more logical and explicit link between criteria that are used in ex ante versus ex post assessments and to relate these to the activities and practices of researchers and research organizations.

The discussion of the potential practitioner implications of the framework also need to address the issue of social responsibility of individual actors and organizations undertaking research activity. In particular, for corporate-led or corporate-supported research, the topic of Corporate Social Responsibility (CSR) should be taken into account. While we could not find specific literature studies investigating the social impact of scientific research involving large organizations and corporations, the relation between science and society is currently at the center of a growing debate that is able to generate new pressures to provide evidence of how and how much science generates social impacts [41].

6.3. Future Research and Concluding Remarks

The increasing societal attention to the strategic relevance of scientific research has amplified national and international systems of research evaluation [1]. Partly driven by global trends, perspectives on research quality have converged around notions of research excellence, which underpin many evaluation regimes [2]. This paper aims primarily to contribute to our understanding of the multiple dimensions of scientific quality. The study bears some limitations. Although the quality evaluation function is mostly a literature-derived framework, further theoretical development and expert analyses are needed to validate the areas and dimensions of evaluation. In addition, a preliminary application of the function to a sample of research units (e.g., research proposals or manuscripts) could provide useful feedback to test and refine the framework. These improvements represent avenues for further research and follow-up analysis to be conducted by the authors as well

as by other members of the research community willing to follow up on the ideas and findings presented here.

Future research could also address the identification of criteria and metrics supporting the evaluation of what represents scientifically "wrong" (or dangerous) research, although if the bibliometrics' performance of the same is positive. This analysis could also permit the strengthening of an editorial or peer-based evaluation of research products before, during or after the review process. Finally, a refinement of the quality dimensions framework could be represented by the ability to "prioritize" items (e.g., using a "stoplight approach") and identify high-relevance, mid-level relevance of less important quality dimensions according to the specific evaluation context or process.

Appendix A. Page for Expert Feedback Collection

Defining and Assessing Quality Dimensions in Scientific Research

#1: The following 77 concepts were obtained from an extensive review of literature on research evaluation, research policy, innovation policy, scientometrics and informetrics. They represent attributes or dimensions (as classified by authors) of quality into scientific research. Please review and suggest any integration of missing attributes (possibly with related literature reference) or advise in order to improve the list.

1. Accessibility
2. Adherence to standards/rules
3. Clarity
4. Clinical significance
5. Coherence
6. Collaboration distance
7. Communicability
8. Community impact
9. Completeness
10. Conformance to ethics
11. Consistency
12. Consumability
13. Contextualization
14. Contributory
15. Cost effectiveness
16. Craftsmanship
17. Credibility
18. Cross-field connections
19. Disinterestedness
20. Dissemination potential
21. Documented sources

22. Economic impact/significance
23. Educational impact
24. Evidence disclosure
25. Expert feedback
26. Feasibility and success probability
27. Generalizability
28. Honesty
29. Impartiality
30. Innovation degree
31. Intellectual significance
32. Interdisciplinarity
33. Internal validity
34. International scope
35. Investigator's expertise
36. Language clarity
37. Methodological soundness
38. Novelty
39. Objectivity
40. Openness
41. Operationalization
42. Organized skepticism
43. Originality
44. Outcomes anticipation
45. Overall research approach
46. Plausibility
47. Political and social significance
48. Practitioner impact
49. Relevance
50. Reliability
51. Reproducibility
52. Research process evaluation
53. Resource attraction
54. Resource management
55. Rigorousness
56. Scholarly exchange and diffusion
57. Scientific significance and value
58. Scope tailoring and focus
59. Searchability
60. Simplicity
61. Social impact/significance
62. Societal relevance and value
63. Stakeholder involvement
64. Stringent argumentation
65. Structure clarity
66. Sustainability
67. Systems view
68. Technological significance
69. Testability
70. Theory impact
71. Thoroughness
72. Transparency
73. Trendsetting and future outline
74. Truth and veridicity
75. Unconventionality

76.   Universalism
77.   Usefulness

The following 5 categories could be used to cluster the 77 concepts above. Please review the list and suggest any amendments in order to enrich/improve the taxonomy:

1.   Quality attributes/dimensions related to Research Vision (i.e., aims, rationale).
2.   Quality attributes/dimensions related to Research Process (i.e., execution, method, conduct);
3.   Quality attributes/dimensions related to Research Description (i.e., formalization, reporting);
4.   Quality attributes/dimensions related to Research Diffusion (i.e., sharing, adoption);
5.   Quality attributes/dimensions related to Research Impact (i.e., effects, influence).

References

1.   Whitley, R.; Glaser, J. *The Changing Governance of the Sciences: The Advent of Research Evaluation Systems*; Springer: Dordrecht, The Netherlands, 2008.
2.   Flink, T.; Peter, T. Excellence and frontier research as travelling concepts in science policymaking. *Minerva* **2018**, *56*, 431–452. [CrossRef]
3.   Sayed, O.H. Critical Treatise on University Ranking Systems. *Open J. Soc. Sci.* **2019**, *7*, 39–51. [CrossRef]
4.   van Vught, F.A.; Ziegele, F. *Design and Testing the Feasibility of a Multidimensional Global University Ranking*; Final Report; European Community, Europe; CHERPA-Network; Cnsortium for Higher Education and Research Performance Assessment: Virtual Network, 2011.
5.   Cole, S. Citations and the evaluation of individual scientists. *Trends Biochem. Sci.* **1989**, *4*, 9–13. [CrossRef]
6.   Gisvold, S.E. Citation analysis and journal impact factors—Is the tail wagging the dog? *Acta Anaesthesiol. Scand.* **1999**, *43*, 971–973. [CrossRef]
7.   Romano, N.C., Jr. Journal self-citation v: Coercive journal self-citation—Manipulations to increase impact factors may do more harm than good in the long run. *Commun. Assoc. Inf. Syst.* **2009**, *25*, 41–56. [CrossRef]
8.   Seglen, P.O. Citation frequency and journal impact: Valid indicators of scientific quality? *J. Intern. Med.* **1991**, *229*, 109–111. [CrossRef] [PubMed]
9.   Seglen, P.O. Citations and journal impact factors: Questionable Indicators of research quality. *Allergy* **1997**, *52*, 1050–1056. [CrossRef]
10. Donovan, C. The qualitative future of research evaluation. *Sci. Public Policy* **2007**, *34*, 585–597. [CrossRef]
11. Hicks, D.; Wouters, P.; Waltman, L.; de Rijcke, S.; Rafols, I. The Leiden Manifesto for research metrics. *Nature* **2015**, *520*, 429–431. [CrossRef]
12. Wilsdon, J.; Belfiore, E.; Campbell, P.; Curry, S.; Hill, S.; Jones, R.; Kain, R.; Kerridge, S.R.; Thelwall, M.; Tinkler, J.; et al. The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessment and Management (HEFCE, 2015). Available online: https://www.researchgate.net/publication/279402178_The_Metric_Tide_Report_of_the_Independent_ Review_of_the_Role_of_Metrics_in_Research_Assessment_and_Management (accessed on 9 February 2022).
13. Leydesdorff, L.; Wouters, F.; Bornman, L. Professional and citizen bibliometrics: Complementarities and ambivalences in the development and use of indicators. *Scientometrics* **2016**, *109*, 2129–2150. [CrossRef]
14. Mårtensson, P.; Fors, U.; Wallin, S.-B.; Zander, U.; Nilsson, G.H. Evaluating research: A multidisciplinary approach to assessing research practice and quality. *Res. Policy* **2016**, *45*, 593–603. [CrossRef]
15. Benedictus, R.; Miedema, F.; Ferguson, M.W. Fewer numbers, better science. *Nature* **2016**, *538*, 453. [CrossRef] [PubMed]
16. Lamont, M.; Fournier, M.; Guetzkow, J.; Mallard, G.; Bernier, R. Evaluating Creative Minds: The Assessment of Originality in Peer Review. In *Knowledge, Communication and Creativity*; Sales, A., Fournier, M., Eds.; SAGE Publications: London, UK, 2007; pp. 166–181.
17. Moore, S.; Neylon, C.; Paul Eve, M.; O'Donnell, D.P.; Pattinson, D. Excellence R Us: University research and the fetishisation of excellence. *Palgrave Commun.* **2017**, *3*, 16105. [CrossRef]
18. Garfield, E. Citation indexes for science: A new dimension in documentation through association of ideas. *Science* **1955**, *122*, 108–111. [CrossRef] [PubMed]
19. Garfield, E. Citation analysis as a tool in journal evaluation. *Science* **1972**, *178*, 471–479. [CrossRef]
20. Garfield, E. Journal impact factor: A brief review. *CMAJ* **1999**, *161*, 977–980.
21. Hirsch, J.E. An index to quantify an individual's scientific research output. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 16569–16572. [CrossRef]
22. Rousseau, R. New developments related to the Hirsch index. *Sci. Focus* **2006**, *1*, 23–25.
23. Egghe, L. Theory and practice of the g-index. *Scientometrics* **2006**, *69*, 131–152. [CrossRef]
24. Alonso, S.; Cabrerizo, F.; Herrera-Viedma, E.; Herrera, F. Hg-index: A new index to characterize the scientific output of researchers based on h- and g-indices. *Scientometrics* **2010**, *82*, 391–400. [CrossRef]

25. Wildgaard, L.; Schneider, J.W.; Larsen, B. A review of the characteristics of 108 author-level bibliometric indicators. *Scientometrics* **2014**, *101*, 125–158. [CrossRef]
26. Waltman, L. A review of the literature on citation impact indicators. *J. Informetr.* **2016**, *10*, 365–391. [CrossRef]
27. Hammarfelt, B.; Rushforth, A.D. Indicators as judgment devices: An empirical study of citizen bibliometrics in research evaluation. *Res. Eval.* **2017**, *26*, 169–180. [CrossRef]
28. Karpik, L. *Valuing the Unique the Economics of Singularities*; Princeton University Press: Princeton, NJ, USA, 2010.
29. Butler, L. Assessing university research: A plea for a balanced approach. *Sci. Public Policy* **2007**, *34*, 565–574. [CrossRef]
30. Moed, H.F. The future of research evaluation rests with an intelligent combination of advanced metrics and transparent peer review. *Sci. Public Policy* **2007**, *34*, 575–583. [CrossRef]
31. Holbrook, J.B.; Barr, K.R.; Brown, K.W. Research Impact: We need negative metrics too. *Nature* **2013**, *497*, 439. [CrossRef]
32. Gibbons, M.; Limoges, C.; Nowotny, H.; Schwartzman, S.; Scott, P.; Trow, M. *The New Production of Knowledge: The Dynamics of Science and Research in Contemporary Societies*; Sage: London, UK, 1994.
33. Gulbrandsen, J.M.; Langfeldt, L. In search of mode 2: The nature of knowledge production in Norway. *Minerva* **2004**, *42*, 237–250. [CrossRef]
34. Albert, M.; Laberge, S.; McGuire, W. Criteria for assessing quality in academic research: The views of biomedical scientists, clinical scientists and social scientists. *High. Educ.* **2012**, *64*, 661–676. [CrossRef]
35. Horbach, S.P.J.M.; Halffman, W. Journal peer review and editorial evaluation: Cautious innovator or sleepy giant? *Minerva* **2020**, *58*, 139–161. [CrossRef]
36. Cronin, B.; Sugimoto, C.R. *Beyond Bibliometrics. Harnessing Multidimensional Indicators of Scholarly Impact*; MIT Press: Cambridge, MA, USA, 2014.
37. De Bellis, N. *Bibliometrics and Citation Analysis: From the Science Citation Index to Cybermetrics*; Scarecrow Press: Lanham, MD, USA, 2009.
38. Prins, A.A.M.; Costas, R.; van Leeuwen, T.N.; Wouters, P.F. Using Google Scholar in research evaluation of humanities and social science programs: A comparison with web of science data. *Res. Eval.* **2016**, *25*, 264–270. [CrossRef]
39. Bornmann, L. Do altmetrics point to the broader impact of research? An overview of benefits and disadvantages of altmetrics. *J. Informetr.* **2014**, *8*, 895–903.
40. Penfield, T.; Baker, M.J.; Scoble, R.; Wykes, M.C. Assessment, evaluations, and definitions of research impact: A review. *Res. Eval.* **2014**, *23*, 21–32. [CrossRef]
41. Reale, E.; Avramov, D.; Canhial, K.; Donovan, C.; Flecha, R.; Holm, P.; Larkin, C.; Lepori, B.; Mosoni-Fried, J.; Oliver, E.; et al. A review of literature on evaluating the scientific, social and political impact of social sciences and humanities research. *Res. Eval. Open Access* **2018**, *27*, 298–308. [CrossRef]
42. Bornmann, L. Measuring the societal impact of research. *EMBO Rep.* **2012**, *13*, 673–676. [CrossRef]
43. Bornmann, L. What is societal Impact of research and how can it be assessed? A literature survey. *J. Am. Soc. Inf. Sci. Technol.* **2013**, *64*, 217–233. [CrossRef]
44. Molas-Gallart, J.; Salter, A.; Patel, P.; Scott, A.; Duran, X. *Measuring Third Stream Activities. Final Report to the Russell Group of Universities*; Science and Technology Policy Research Unit (SPRU): Brighton, UK, 2002.
45. Van der Meulen, B.; Rip, A. Evaluation of societal quality of public sector research in the Netherlands. *Res. Eval.* **2000**, *9*, 11–25. [CrossRef]
46. DEST–Department of Education Science and Training. *Research Quality Framework: Assessing the Quality and Impact of Research in Australia*; Commonwealth of Australia: Canberra, Australia, 2005.
47. Hemlin, S. Utility evaluation of academic research: Six basic propositions. *Res. Eval.* **1998**, *7*, 159–165. [CrossRef]
48. McNie, E.C.; Parris, A.; Sarewitz, D. Improving the public value of science: A typology to inform discussion, design and implementation of research. *Res. Policy* **2016**, *45*, 884–895. [CrossRef]
49. Bozeman, B.; Sarewitz, D. Public value mapping and science policy evaluation. *Minerva* **2011**, *49*, 1–23. [CrossRef]
50. ERiC. *Evaluating the Societal Relevance of Academic Research: A Guide*; Rathenau Institute: The Hague, The Netherlands, 2010.
51. Holbrook, J.B.; Frodeman, R. Peer review and the ex-ante assessment of societal impacts. *Res. Eval.* **2011**, *20*, 239–246. [CrossRef]
52. Morton, S. Progressing research impact assessment: A contributions' approach. *Res. Eval.* **2015**, *24*, 405–419. [CrossRef]
53. Holbrook, J.B.; Hrotic, S. Blue skies, impacts, and peer review. *Roars Trans. J. Res. Policy Eval. (RT)* **2013**, *1*, 1–24.
54. ERC—European Research Council. *Qualitative Evaluation of Completed Projects Funded by the European Research Council*; European Commission: Brussels, Belgium, 2016.
55. Samuel, G.N.; Derrick, G. Societal Impact evaluation: Exploring evaluator perceptions of the characterization of impact under the ref2014. *Res. Eval.* **2015**, *24*, 229–241. [CrossRef]
56. Lamont, M. *How Professors Think: Inside the Curious World of Academic Judgment*; Harvard University Press: Cambridge, MA, USA, 2009.
57. Bazeley, P. Conceptualising research performance. *Stud. High. Educ.* **2010**, *35*, 889–903. [CrossRef]
58. Llonk, J.; Casablancas-Segura, C.; Alarcón-del-Amo, M.C. Stakeholder orientation in public universities: A conceptual discussion and a scale development. *Span. J. Mark. ESIC* **2016**, *20*, 41–57. [CrossRef]
59. Bras-Amorós, M.; Domingo-Ferrer, J.; Torra, V. A bibliometric index based on the collaboration distance between cited and citing authors. *J. Informetr.* **2011**, *5*, 248–264. [CrossRef]

60. Lawani, S.M. Some bibliometric correlates of quality in scientific research. *Scientometrics* **1984**, *9*, 13–25. [CrossRef]
61. Hemlin, S. *Quality in Science. Researchers' Conceptions and Judgements*; University of Gothenburg, Department of Psychology: Gothenburg, Sweden, 1991.
62. Gulbrandsen, J.M. *Research Quality and Organisational Factors: An Investigation of the Relationship*; Norwegian University of Science and Technology: Trondheim, Sweden, 2000.
63. Keen, P.G.W. Relevance and rigor in information systems research: Improving quality, confidence, cohesion and impact. In *Information Systems Research: Contemporary Approaches and Emergent Traditions*; Klein, H.K., Nissen, H.-E., Hirschheim, R., Eds.; IFIP Elsevier Science: Philadelphia, PA, USA, 1991; pp. 27–49.
64. Maxwell, J.A. *Qualitative Research Design: An Interactive Approach*; Sage Publications: Thousand Oaks, CA, USA, 1996.
65. Amin, A.; Roberts, J. Knowing in action: Beyond communities of practice. *Res. Policy* **2008**, *37*, 353–369. [CrossRef]
66. Langfeldt LNedeva, M.; Sörlin, S.; Thomas, D.A. Co-existing notions of research quality: A framework to study context-specific understandings of good research. *Minerva* **2020**, *58*, 115–137. [CrossRef]
67. Baldi, S. Normative versus social constructivist processes in the allocation of citations: A network-analytic model. *Am. Sociol. Rev.* **2008**, *63*, 829–846. [CrossRef]
68. Silverman, D. *Interpreting Qualitative Data: Methods for Analysing Talk, Text and Interaction*; Sage Publications: London, UK, 1993.
69. Alborz, A.; McNally, R. Developing methods for systematic reviewing in health services delivery and organisation: An example from a review of access to health care for people with learning disabilities. Part Evaluation of the literature—A practical guide. *Health Inf. Lib. J.* **2004**, *21*, 227–236. [CrossRef] [PubMed]
70. Lee, C.J. Commensuration bias in peer review. *Philos. Sci.* **2015**, *82*, 1272–1283. [CrossRef]
71. Polanyi, M. The republic of science: Its political and economic theory Minerva, I(1) (1962), 54–73. *Minerva* **2000**, *38*, 1–32. [CrossRef]
72. Buchholz, K. Criteria for the analysis of scientific quality. *Scientometrics* **1995**, *32*, 195–218. [CrossRef]
73. Hug, S.E.; Ochsner, M.; Daniel, H.-D. Criteria for assessing research quality in the humanities: A Delphi study among scholars of English literature, German literature and art history. *Res. Eval.* **2013**, *22*, 369–383. [CrossRef]
74. Lahtinen, E.; Koskinen-Ollonqvist, P.; Rouvinen-Wilenius, P.; Tuominen, P.; Mittelmark, M.B. The development of quality criteria for research: A Finnish approach. *Health Promot. Int.* **2005**, *20*, 306–315. [CrossRef]
75. External Research Assessment (ERA). *External Research Assessment*; Karolinska Institutet: Stockholm, Sweden, 2010.
76. Hemlin, S.; Montgomery, H. Scientists' conceptions of scientific quality: An interview study. *Sci. Stud.* **1990**, *3*, 73–81. [CrossRef]
77. Hemlin, S.; Niemenmaa, P.; Montgomery, H. Quality criteria in evaluations: Peer reviews of grant applications in psychology. *Sci. Stud.* **1995**, *8*, 44–52. [CrossRef]
78. Weinberg, A.M. Criteria for scientific choice: Minerva, I (2), (1962), 158–171. *Minerva* **2000**, *38*, 255–266. [CrossRef]
79. Shipman, M. *The Limitations of Social Research*; Longman: London, UK, 1982.
80. Gummesson, E. *Qualitative Methods in Management Research*; Sage Publications: Newbury Park, CA, USA, 1991.
81. Ochsner, M.; Hug, S.E.; Daniel, H.-D. Four types of research in the humanities: Setting the stage for research quality criteria in the humanities. *Res. Eval.* **2012**, *22*, 79–92. [CrossRef]
82. Tranøy, K.E. *Science–Social Power and Way of Life*; Universitetsforlaget: Oslo, Norway, 1986.
83. Guthrie, S.; Wamae, W.; Diepeveen, S.; Wooding, S.; Grant, J. *Measuring Research: A Guide to Research Evaluation Frameworks and Tools*; Rand: Santa Monica, CA, USA, 2013.
84. De Rijcke, S.; Wouters, P.F.; Rushforth, A.D.; Franssen, T.P.; Hammarfelt, B. Evaluation practices and effects of indicator use: A literature review. *Res. Eval.* **2016**, *25*, 161–169. [CrossRef]
85. Waltman, L. Responsible Metrics: One Size Doesn't Fit All, CWTS—Centre for Science and Technology Studies, CWTS Blog article. Available online: https://www.cwts.nl (accessed on 31 August 2017).
86. Kaplan, S.; Norton, D.P. *The Balanced Scorecard*; HBS Press: Boston, MA, USA, 1992.