Advances in the Evaluation of Scientific and Technological Talent: An Analysis Based on Literature and Policy Documents from 2010-2021

Authors: Feng Guohe, Chen Chenyan, Feng Guohe

Date: 2023-01-17T00:00:00+00:00

Abstract

Scientific and technological talent constitutes a special category of human resources—highly skilled individuals who participate directly or indirectly in scientific research and innovation activities. The capacity of this group is directly related to scientific and technological innovation and development; therefore, the evaluation of scientific and technological talent is of paramount importance. Optimizing the evaluation system for scientific and technological talent and enhancing the effectiveness of such evaluations are critical approaches for revitalizing the talent pool and improving the overall quality of the scientific and technological workforce. Employing web survey methodology and content analysis, this study systematically examines literature and policy documents related to scientific and technological talent evaluation from 2010 to 2021, exploring the evolution of evaluation policies and methods, shifts in evaluation philosophy, emerging trends, and existing challenges in scientific and technological talent evaluation. The research findings indicate that as the evaluation system continues to mature, evaluation outcomes have transitioned from being performance-oriented to being innovation- and actual achievement-oriented; evaluation methods have shifted from singular approaches to multi-dimensional, comprehensive frameworks that integrate qualitative and quantitative techniques; and evaluation trends are characterized by new developments and paradigms centered on the core principles of data-driven approaches, open collaboration, representative achievements, and responsible evaluation.

Full Text

Progress in Science and Technology Talent Evaluation: An Analysis Based on Literature and Policy Texts from 2010-2021

FENG Guohe, CHEN Chenyan
School of Economics and Management, South China Normal University, Guangzhou 510006, China

Abstract

[**Purpose/Significance**] As a special category of human resources, science and technology (S&T) talents are highly skilled individuals who directly or indirectly participate in scientific research and innovation activities. The strength of this group directly affects technological innovation and development, making the evaluation of S&T talents particularly crucial. Optimizing the evaluation system and strengthening its effectiveness are key methods for revitalizing talent pools and enhancing the overall quality of S&T talent teams. [**Method/Process**] Based on web surveys and content analysis, this study systematically analyzes literature and policy texts related to S&T talent evaluation from 2010-2021, exploring the evolution of evaluation policies and methods, changes in evaluation concepts, emerging trends, and existing problems. [**Result/Conclusion**] The research reveals that as the S&T talent evaluation system continues to improve, evaluation results have shifted from performance-oriented to innovation and actual achievement-oriented approaches; evaluation methods have transformed from single-dimensional to multi-dimensional, comprehensive approaches combining qualitative and quantitative techniques; and new trends are emerging centered on data-driven, open collaboration, representative achievements, and responsible evaluation.

Keywords: scientific and technological talents; evaluation theory; evaluation elements; evaluation trends
Classification Number: G350

Since the reform and opening up, China has adhered to the fundamental policy that "science and technology are primary productive forces." Over four decades of development, remarkable achievements have been made in economy, science and technology, and culture. In the new era, China' s economic focus has shifted from pursuing GDP growth to industrial restructuring and high-quality development, with scientific and technological innovation at the core of this transformation [1]. In 2021, President Xi Jinping emphasized that "we must steadfastly take scientific and technological self-reliance and self-improvement as the strategic support for national development." In this process of national innovation development, S&T talents represent a special human resource and the most critical asset in technological innovation. As highly skilled individuals directly involved in research and innovation activities, their scientific capabilities

and innovative vitality are closely linked to high-quality research output. For research teams, S&T talent evaluation helps identify candidates suitable for research positions; for the talents themselves, evaluation results are often tied to material rewards, promotion opportunities, and honors. Scientific, rational, and fair evaluation enhances the motivation of S&T talents and correctly guides the value orientation of research work.

To guide and standardize S&T talent evaluation, the Chinese government has issued a series of policy documents. The 2006 *National Medium- and Long-Term Program for Scientific and Technological Development (2006-2020)* explicitly recognized the important role of S&T talents in research innovation and first proposed cultivating a high-level talent team while tackling technical challenges. The June 2010 *National Medium- and Long-Term Talent Development Plan (2010-2020)* identified research performance as a crucial criterion and emphasized overall talent quality cultivation and innovative talent mechanisms. In February 2018, the General Office of the CPC Central Committee and the State Council issued the *Guiding Opinions on Deepening the Reform of Talent Evaluation Mechanisms by Category* (hereinafter referred to as the *Categorical Evaluation Guiding Opinions*), which clearly stated that evaluation standards should be set according to different talent characteristics, with "moral character as the primary content" of evaluation, and aimed to curb the "three-only" tendency (only academic credentials, only seniority, and only papers). Subsequently, various ministries issued documents to "break the four onlys" and "break the five onlys," profoundly influencing S&T talent evaluation.

This paper conducts web surveys and content analysis of literature and policy texts to explore the evolution of S&T talent evaluation from multiple dimensions, identify problems, and propose recommendations. The structure is as follows: Section 1 introduces the background; Section 2 analyzes concepts, categories, policy evolution, supporting theories, evaluation effectiveness, and methods; Section 3 examines existing problems; and Section 4 discusses emerging trends.

2 Multi-dimensional Analysis of S&T Talent Evaluation

Based on literature and policy texts, this section explores changes in S&T talent evaluation from multiple dimensions to promote its development in a scientific, correct, rational, and fair direction.

2.1.1 Concepts and Categories of S&T Talent

To date, academia has not reached a unified definition of S&T talent. As early as 2004, Du Qian [2] categorized perceptions of S&T talent into three types: scientific and technological elites, highly educated individuals in relevant fields, and S&T activity personnel. In 2011, the Ministry of Science and Technology's *National Medium- and Long-Term Plan for S&T Talent Development (2010-2020)* defined S&T talent as "workers with certain professional knowledge or

specialized skills who engage in creative scientific and technological activities and contribute to S&T endeavors and economic and social development." Scientific and rational classification of S&T talent is the prerequisite for establishing a categorical evaluation system. The author summarizes domestic research on S&T talent classification in Table 1 .

The connotation of S&T talent continues to adjust and expand with social development and policy changes. The above literature classifies talent based on disciplinary categories and technical characteristics, providing a basis for the current categorical evaluation system. Talent classification is the top-level design and important guarantee for implementing tiered management and resource allocation. The 2017 *13th Five-Year Plan for National S&T Talent Development* explicitly proposed implementing management methods such as "categorical evaluation," "categorical policy support," and "categorical construction of S&T management training systems," fully demonstrating that S&T talent classification is the prerequisite for precise policy implementation and rewards/punishments, and an important factor affecting policy implementation and effectiveness evaluation. Talent classification clarifies the work positioning of different positions and categories of S&T talent, encourages each to perform their duties, and provides a basis for building a rational S&T talent evaluation framework.

2.1.2 Related Policies on S&T Talent Evaluation

To ensure scientific and comprehensive policy retrieval results, the author selected high-frequency terms, main themes, and expanded terms from S&T talent evaluation literature in CNKI, including "scientific and technological innovation," "S&T talent," "S&T talent rating," and "talent evaluation" to conduct full-text policy searches on the Central People' s Government portal, PKULaw database, and other departmental websites. Considering the timeliness, effectiveness level, and applicability of policy documents, the author excluded local regulations outside the research timeframe, expired documents, and those with limited influence, selecting 18 policy documents as shown in Table 2 .

Since the release of the *2010 Talent Development Outline*, China has continuously issued S&T support policies, consistently prioritizing technological innovation in overall development with increasing investment, aiming to make technology a lever for social productivity transformation and apply research results to actual production and people' s livelihood. As early as 1996, China promulgated the *Law of the People' s Republic of China on Promoting the Transformation of Scientific and Technological Achievements* (hereinafter referred to as the *Achievement Transformation Law*), which was revised in 2015 to optimize intellectual property protection mechanisms, safeguard researchers' knowledge rights, and accelerate achievement transformation. Although the revised law demonstrates increasing emphasis on cultivating innovative S&T talent and transforming achievements, the transformation rate has not met expectations due to structural imbalances in policy tool usage [8] and information disadvan-

tages for technology holders in the market [9]. According to the *China Innovation Index Study* released by the National Bureau of Statistics in October 2021, S&T talent has made tremendous progress in both the quantity and quality of research output, but the transformation rate remains low; innovation achievements are inseparable from economic and industrial development. Reviewing these documents reveals that S&T talent evaluation increasingly emphasizes the importance and necessity of market application, requiring not only market demand orientation but also close integration between achievement transformation and talent evaluation systems, with transformation rates listed as key evaluation indicators. This indicates that China' s current S&T talent evaluation system places greater emphasis on actual research performance and comprehensive contributions, moving away from blindly pursuing high growth in innovation output to focus on innovation effectiveness and application.

2.2 Supporting Theories: Evolution from General to Refined Frameworks

S&T talent evaluation is not only a process of recognizing and affirming researchers but also reflects societal and research teams' internal expectations of them. Theoretical analysis and discussion address why, what, and how to evaluate, revealing several foundational theories:

2.2.1 Individual Difference Theory Before engaging in research activities, all S&T talents share a common identity— "social beings." Different researchers have varying physiological conditions, personality traits, growth environments, and work experiences, which play crucial roles in shaping their professional identities. One function of talent evaluation is to establish scientifically sound methods that link individual comprehensive qualities with long-term team planning, motivating S&T talents to align with team goals. As society develops, external perceptions of S&T talents extend beyond research achievements, incorporating theories from social psychology and economics into evaluation systems. If elements inhibiting productivity are identified, research teams can take timely, targeted measures to revitalize motivation, achieving a win-win situation for both individuals and teams.

The renowned psychologist McGregor used "Theory X-Y" [10] to explain how reward and punishment systems based on performance evaluation help mobilize employee productivity and creativity. Theory X-Y introduced performance indicators into S&T talent evaluation, becoming an important basis for assessment. In 1973, psychologist McClelland first proposed the concept of "competence" [11] and subsequently developed competency models based on different competency combinations. Competency models are widely applied in human resource management [12] and S&T talent evaluation indicator construction [13]. The iceberg model divides personal qualities into explicit competencies (skills and knowledge) and implicit competencies (personal character, work motivation, self-evaluation, and values) [11], providing a theoretical foundation for multi-level,

multi-angle evaluation. In 1975, McClelland proposed the concept of "psychological empowerment" [14], referring to employees' self-efficacy cognition under management practices. The mediating role of psychological empowerment is widely applied in evaluating employee competency [15], professional identity [16], and psychological states [17]. Integrating these theories with talent evaluation systems provides new approaches for improving accuracy, effectiveness, and comprehensiveness. Measuring S&T talent productivity solely by research output quantity yields crude results that fail to capture key information about research processes and comprehensive qualities, reducing the credibility of outcome control and weakening effective incentives.

The 1965 publication *Organizational Psychology* reinterpreted Mayo' s social person hypothesis and proposed the complex person hypothesis [18]. According to this hypothesis, although each researcher possesses independent judgment and complex emotions, individual characteristics and patterns can still be extracted from multi-dimensional indicators such as living environment and work status. Recent research based on these hypotheses has explored factors affecting employee innovation capability and how to enhance innovation drive [19]. Studies show that factors influencing researchers' innovation capacity and enthusiasm include material rewards [20], emotional incentives [21], management systems [22], individual abilities [23], and cooperative-competitive environments [24]. Human essence is the sum of social relations, and changes in social relations directly or indirectly affect S&T talents. Therefore, when building evaluation systems, talent groups should be examined with a dynamic developmental perspective rather than implementing "one-size-fits-all" policies. When formulating talent incentives, besides rational distribution of material rewards and promotion opportunities, value orientation should be emphasized, comprehensively examining S&T talents' comprehensive abilities and moral standards while creating a positive organizational culture. Only through synergistic development of sound management systems and correct research values can overall research output quality be improved.

2.2.2 Fuzzy Thinking Fuzzy thinking involves abstracting and delineating boundaries of concrete phenomena to achieve understanding. In talent evaluation, it manifests as: evaluators' cognitive fuzziness, the developmental fuzziness of evaluation objects' achievements, and fuzziness in the evaluation process. Research achievements often undergo peer review, where experts' limitations in research fields, professional ethics, and personal preferences prevent complete objectivity, making review opinions products of fuzzy cognition. Achievement recognition should be tested by markets and users, but since research is long-term with uncertain hypotheses, evaluated achievements are often parts of long-term projects or completed projects with insufficient practical feedback time. This creates uncertainty in achievement transformation and developmental fuzziness. Additionally, review teams experience a "fuzzy-precise-fuzzy" [25] thinking process when analyzing submitted materials: initially "fuzzily" understanding materials according to their own thinking patterns, then extracting

relatively "precise" contributions, and finally giving "fuzzy" definitions through scoring, comments, or voting. Expert evaluation and peer review are common applications of fuzzy thinking, providing comprehensive understanding of evaluation objects' work and life with considerable credibility as qualitative methods. However, these methods are highly subjective and fuzzy. To improve accuracy, scholars have introduced fuzzy mathematics and fuzzy neural networks [26] into S&T talent evaluation, making fuzzy thinking more scientific and rational.

2.2.3 Measurability of Phenomena Material objects possess objective reality and can be measured. S&T talents also have objective reality, with varying social contributions to research achievements influenced by their abilities, intelligence, and character, making objective research performance measurable [27]. Measuring S&T talent effectiveness essentially depicts development patterns, reflecting dynamic research processes through static results to provide quantifiable work performance for organizations and feedback for talents. S&T talent evaluation must follow the principle of measurement consistency to ensure fairness and timely feedback. As early as 2002, Qiu Junping [28] summarized basic talent evaluation elements into six indicators: moral character, scientific culture, physical and mental health, professional competence, innovation, and collaboration. Based on the ultimate goals of "strengthening the nation through science and technology" and "benefiting the people through science and technology," and considering different position natures across talent categories, scholars have incorporated achievement transformation and social practice into basic evaluation elements [29]. Chen Jianxin proposed that talent evaluation should be diversified and socialized [30]. With the advent of big data, scholars have introduced quantitative methods such as search algorithms [31], BP neural networks [32], and grey relational analysis [33] into talent evaluation to improve precision and support management decision-making.

2.3 Evaluation Results: Shifting from Performance-Oriented to Innovation, Quality, Effectiveness, and Contribution-Oriented

S&T talents are indispensable actors in research innovation, mastering cutting-edge research dynamics for nations and regions. Studies show that S&T human resource investment positively correlates with national innovation development [34], while the original evaluation mechanisms were insufficient for talent motivation and research transformation needs [35]. Therefore, while increasing research investment, organizations must also address S&T talent needs and achievement transformation. Scientific evaluation systems positively influence researcher motivation and innovation vitality, serving a value orientation function.

Evaluation results directly affect S&T talents' treatment and promotion opportunities. Chronologically reviewing national policies on talent team categorical management reveals that since reform and opening up, S&T talent evaluation has experienced four periods: exploration (1978-1991), formation (1992-2001), development (2002-2017), and deepening (2018-present), with different value

orientations in each period.

Taking 2006 as a policy turning point, earlier policies covered broad talent scopes including innovation talents, industry skilled talents, and vocational technical talents without precisely proposing management methods for improving S&T talent evaluation systems. However, the shift from performance-oriented to innovation and quality-oriented evaluation was already emerging. The 2004 *State Administration of Radio, Film and Television' s Opinions on Further Strengthening Radio, Film and Television Talent Work* stated that evaluation mechanisms should be ability and performance-oriented. In 2010, the State Intellectual Property Office announced that evaluation should be based on moral character, ability, and performance. The 2006 *National 11th Five-Year Plan for Science and Technology Development* proposed establishing an ability-oriented S&T talent evaluation mechanism, marking the first time S&T talent appeared as a specific evaluation category in policy texts. The 2011 *Talent Development Plan* first proposed establishing a categorical evaluation system for S&T talent. The 2012 *Opinions on Deepening Science and Technology System Reform and Accelerating National Innovation System Construction* indicated shifting evaluation orientation to research capability and innovative achievements. In 2018, the Ministry of Education issued the *Notice on Breaking the "Four Onlys"* and *Notice on Breaking the "Five Onlys,"* demonstrating government determination for bold reform. As the scientific community sought to break the "five onlys," Liu Yidong [36] advocated using a cutting-edge scholar responsibility system to select academic leaders, shifting evaluation focus from titles to cutting-edge achievements and peer recognition to identify innovative high-quality talent.

Abandoning "five onlys" thinking is essential to correct the misalignment between evaluation results and actual innovation performance. Professional titles, honors, and awards should be byproducts of innovation incentive systems, not ends pursued through consuming national resources to follow research trends, which violates policy intentions and hinders high-level talent cultivation.

Research shows that S&T system reform policies positively moderate S&T talent innovation investment [37], serving as a baton for talent work priorities. Reviewing policy evolution reveals China' s talent evaluation is shifting from performance-oriented to comprehensive contribution-oriented, transforming from standards emphasizing paper quantity, titles, and awards to multi-angle criteria including ideological quality and academic capability.

To meet requirements for accurately understanding talent status and scientific selection, evaluation dimensions have shifted from basic similarity and overall convergence to multi-dimensional, multi-level approaches. Traditional systems focused on academic contributions, social benefits, and economic benefits [38], emphasizing S&T contributions while neglecting self-evaluation [39]. From the talent perspective, although current policies emphasize innovation development, they remain largely oriented toward improving research capability with insufficient incentives [40].

2.4 Evaluation Methods: Transitioning from Purely Quantitative or Qualitative to Combined Approaches

Technological innovation is essential for social progress, and establishing a scientific talent evaluation system helps stimulate work enthusiasm, tap team innovation potential, and improve talent quality. Since scientific achievements are typically completed through team collaboration, appropriate intervention in evaluating S&T talents' professional capabilities enhances team efficiency. Literature review shows bibliometric analysis and peer review are the most fundamental and commonly used methods.

China' s S&T talent evaluation once fell into single-indicator approaches. In the 1970s, bibliometrics emerged in talent evaluation, and by the 1980s it was considered a universal method for research performance assessment [41]. During the early development of evaluation mechanisms, bibliometric indicators became virtually the sole criterion [42], determining talents' career advancement. Traditional bibliometric indicators include publication quantity, journal impact factor, citation frequency, and H-index [43]. Impact factor measures the ratio of total citations to papers published in a journal over two years, evaluating journal influence but not directly measuring individual achievement quality [44]. Citation frequency counts total references to specific literature in databases, with higher frequencies indicating greater field importance and indirectly reflecting content quality. However, database preferences make sole reliance on citations incomplete and unscientific for evaluating individual contributions or team rankings [45]. The H-index measures cumulative research impact and relevance [46] but has limitations in cross-disciplinary comparisons and quality assessment, and only increases over time, failing to describe career-stage variations scientifically. Thus, quantitative analysis using only a few indicators cannot reasonably evaluate S&T talent.

China' s WTO accession in December 2001 laid a foundation for globalization in economy and culture, providing broader international perspectives for scholars. Research from the U.S. National Science Foundation showed that single bibliometric indicators are unsuitable for measuring individual contributions [48]. Following the 2006 *National Medium- and Long-Term Program for Scientific and Technological Development*, S&T talents' important status became more prominent, and scholars realized that prioritizing publications could mislead work orientation. While bibliometrics offers objectivity, repeatability, authority, and convenience, it cannot assess "spillover benefits" or the authenticity and practicality of achievements [49]. Bibliometrics, being output-oriented, has long served as an important research performance basis but only reflects input-output efficiency without demonstrating research timeliness or market recognition. Consequently, many scholars refuted "single bibliometric indicator theory" and expanded evaluation indicators, arguing that research evaluation should address both "quality" and "quantity," with measurement being only one part, and the key being finding efficient precision methods within existing indicators [50].

Peer review is a qualitative research quality evaluation method. Britain first adopted it three centuries ago, and it was introduced to China in the late 1970s [51] for project management and cadre assessment in medical and commercial technology fields. In the early 1980s, it was applied to doctoral dissertation evaluation and later to research project execution and assessment at all levels. The 2011 *Talent Development Plan* explicitly introduced international peer review standards for basic research talent evaluation. Since then, peer review has been widely applied in achievement evaluation across academic journals [52] and scientific funds [53], providing authoritative results. However, with explosive growth in research output, traditional peer review has revealed limitations including low efficiency [54] and insufficient transparency in blind review systems [55]. To achieve greater fairness and efficiency, peer review has continuously transformed: from the traditional "submission-peer review-publication" model to preprint-based direct peer evaluation; from vague review standards to non-selective peer review for greater objectivity; and from blind review to open peer review for more standardized and fair processes [56]. To establish a fairer and more transparent evaluation system, experts continuously 赋予 new connotations to evaluation purposes, refine application scenarios, and expand evaluation factors.

As evaluation systems mature, evaluation object boundaries have broadened to include not only academic achievements but also academic influence and ideological quality. The 2016 *Opinions on Deepening Talent Development System and Mechanism Reform* stated that evaluation should highlight moral character, individual capability, and research performance. Since different indicators have different natures and content, relying solely on quantitative or qualitative methods cannot meet comprehensive evaluation needs. Consequently, scholars have introduced combined quantitative-qualitative methods into evaluation model construction, including typical approaches such as Analytic Hierarchy Process [57], fuzzy comprehensive evaluation [58], Delphi method [59], and factor analysis [60].

Current evaluation systems generally employ multiple methods, considering S&T talents' social and developmental nature. Evaluation typically combines subjective and objective assessment, integrating the scientific nature of quantitative analysis with the logic and experience of qualitative analysis to multidimensionally characterize contributions. Indicators are adjusted according to talent characteristics and position requirements, with rational use of both methods to provide a fair and reliable evaluation environment.

3 Existing Problems in S&T Talent Evaluation

3.1 Problems with Evaluation Subjects

Evaluation subjects are those exercising evaluation authority over academic achievements, primarily including superior/subordinate organizations, evaluation agencies, domain experts, and the public. Traditional research evalua-

tion systems are based on the principal-agent theory proposed by American economists Berle and Means, advocating separation of quality monitoring rights from research execution rights [61]. During evaluation, a "public-government-evaluation agency/expert panel" principal-agent chain forms, delegating evaluation to professional agencies or panels. As the state increasingly emphasizes evaluation system standardization and rationality, problems within evaluation subjects have attracted attention. The 2018 *Opinions on Deepening Project Review, Talent Evaluation, and Institutional Assessment Reform* emphasized the "principal position of employing units." Commissioned evaluation agencies often lack understanding of evaluation objects' workflows and scopes, making comprehensive judgments difficult and causing reliance on traditional bibliometric indicators that easily fall into "rational tool supremacy" evaluation modes.

Delegating evaluation authority to employing units strengthens their research autonomy. Units can conduct quality control from the source, formulate regulations and evaluation systems aligned with team development, and implement categorical and tiered evaluation based on position responsibilities and work content. This helps resolve structural contradictions between actual contributions and incomplete third-party results, stimulating internal innovation drive. Based on responsible metrics and evaluation concepts, evaluation subjects should follow objective realities, fully understand workflows, content, object quality, and contributions, and accurately characterize evaluation objects with objectivity and fairness in the spirit of academic public service.

3.2 Problems with Evaluation Objects

Current discussions on S&T talent evaluation focus primarily on methods and indicators, with limited research on talent classification [62]. As disciplinary research and knowledge application levels develop, knowledge production modes have shifted from single-disciplinary to market-demand-oriented research emphasizing coordinated development of resources, environment, and research subjects, with disciplines rising spirally through intersection [63]. Traditional extensive evaluation models no longer suit current innovation situations.

The 2018 *Guiding Opinions on Deepening the Reform of Talent Evaluation Mechanisms by Category* clearly stated that "a scientific talent categorical evaluation system should be improved based on professional attributes and position requirements." Talent classification is crucial for correctly understanding disciplinary characteristics, talent status, and position functions, and is the prerequisite for scientific and refined evaluation. In practice, disciplines like medicine [64] and agriculture [65] that integrate practice and research often have blurred position boundaries and require "versatility," leading to researchers holding multiple positions. This increases workload and causes mismatches between work and position requirements, as well as unhealthy competition in task allocation. A scientific categorical evaluation system is the premise for precise talent positioning. Only through proper classification and clear responsibilities and goals can division of labor be clarified, allowing professionals to do professional work,

strengthening internal management control, avoiding responsibility-shifting, ensuring evaluation accuracy and comprehensiveness, motivating teamwork, enhancing cohesion, and creating a positive research atmosphere. Following categorical and tiered principles enables precise evaluation, reflecting responsible evaluation concepts. Only through precise evaluation can S&T talent enthusiasm be truly stimulated, maximizing their value for greater national contributions.

3.3 Problems with Evaluation Methods

Traditional evaluation methods primarily include peer review and bibliometrics. (1) **Peer Review.** Peer review is currently the most commonly used qualitative method for academic achievement evaluation worldwide. Jiang Xiaohua studied appointment and assessment systems at 10 world-class universities, finding that peer review was prioritized in performance evaluation [66]. With characteristics of universalism, professional authority, academic community, and economy [67], peer review has become the "gatekeeper" of academic access, establishing an unshakable position in research evaluation. However, as research scope expands and interdisciplinary integration deepens, peer review has revealed drawbacks. First, interdisciplinary phenomena increase difficulty in selecting reviewers, raise communication costs, and reduce process control [68]. Second, academic frameworks cannot completely separate from personal networks. As noted in *How Professors Think: Inside the Curious World of Academic Judgment*, scholars' research fields intersect with social relationships [69]. Peer review originates from academic autonomy, entrusting specific fields to experts, but scholars in specific fields often have varying relationships. Academic conferences provide platforms for exchange, and even without real-life connections, vertical research fields require scholars to understand each other' s work and frontiers. These phenomena can compromise transparency and scientific rigor. Third, peer review results are susceptible to the "Matthew effect" in academia. Academic leaders often obtain richer resources, producing more and better results and gaining more honors. When evaluating high-reputation talents or teams, experts may fall into authority worship, affecting fairness. Young researchers with less impressive credentials struggle to break through established authority; if evaluation subjects are not objective and fair, young researchers' achievements may not be recognized [70], affecting their enthusiasm and efficiency. On August 8, the Ministry of Science and Technology issued the *Notice on Special Action to Reduce Burdens on Young Researchers*, explicitly addressing problems including limited opportunities, narrow growth channels, frequent evaluations, and heavy administrative burdens, ensuring young researchers can focus on research [71]. These policies compel us to break existing evaluation modes, implement layered and categorical targeted evaluation, unburden young talent, and allow them to concentrate on tackling difficult problems and pursuing original innovation.

(2) **Bibliometrics.** Bibliometrics quantifies research achievements, academic influence, and innovation capability. Common indicators include publica-

tion quantity, citation frequency, Hirsch index, and journal impact factor. Quantitative indicators offer good compatibility and interpretability, intuitively expressing research output efficiency and converting production materials into clear, standardized structured data, reducing evaluation costs and improving management decision-making efficiency. However, problems remain. First, excessive use of measurement indicators leads to blind pursuit of instrumental rationality. While measurement indicators' extensibility makes them important across disciplines and positions, simplifying evaluation processes greatly increases misuse probability. The 2018 *Notice on Cleaning Up "Only Papers, Only Titles, Only Academic Credentials, Only Awards"* emphasized the necessity of "breaking the five onlys," advocating focus on research itself, introducing multi-dimensional indicators, reducing utilitarian drivers in publications, and creating a healthy research environment [72]. Second, quantitative indicators easily cause "emphasizing quantity over quality" and "evaluating papers by journals." Bibliometrics originally provided convenient, repeatable evaluation frameworks for similar fields, but overemphasis on quantitative indicators encourages undesirable behaviors like blindly applying for projects and padding papers to increase publication numbers and beautify resumes. If S&T talents approach projects with such attitudes, research quality suffers, wasting public resources and eroding the authority and fairness of China' s research evaluation system, severely undermining overall innovation vitality. The 2020 *Fifth Round Discipline Evaluation Work Plan* explicitly stated that diagnostic functions should be highlighted, emphasizing the importance of research quality, academic contribution, and innovation capability, while preventing behaviors and value orientations detrimental to high-quality output. Third, quantitative indicators struggle to measure social contributions. Current discipline evaluation methods typically use publication numbers in high-reputation journals and national/provincial project/award numbers as evaluation standards [73]. To meet these standards, talents and institutions may target these indicators, blindly pursuing academic hotspots and short-term gains instead of conducting serious research. Worse, many projects lower quality and increase publication numbers near evaluation deadlines, abandoning previous research afterward to pursue new hotspots and projects.

The purpose of research evaluation is to motivate S&T talents to create more value, solve practical problems, and contribute to society. Research is not accomplished overnight; the process from problem identification to solution implementation and application is lengthy, requiring tremendous effort and perseverance. If researchers cannot diligently cultivate their fields, their output will be superficial. Existing methods are not universally applicable; different types and positions require different approaches. For example, basic research values original innovation, where high-quality papers are important carriers, making them feasible representative works. Applied basic research emphasizes transformation and application, where innovative patents can serve as representative

works. Different quantitative indicators require different evaluation methods in specific assessments, reflecting responsible evaluation.

3.4 Problems with Evaluation Process

The evaluation process is a critical step where evaluation subjects rank contributions based on research performance, characterized by fuzziness and empiricism. Problems include: First, **uncertainty in disciplinary affiliation.** In an open academic environment, breaking disciplinary barriers and interdisciplinary interaction have become unstoppable trends. Old boundaries are blurring, and transdisciplinary concepts are spreading, pressuring existing systems. When declaring achievements, talents report according to their fields; for multidisciplinary collaborative achievements, declaration is negotiated by institutions based primarily on the leader's specialty. This "one-size-fits-all" approach causes mismatches between achievements and declared directions, and recommended reviewers may not understand interdisciplinary results, distorting evaluation outcomes. Second, **insufficient process data.** Process data reflects daily work content and achievements. Innovation is a long-term, dynamic process with varying "bottleneck" problem difficulty and disciplinary accumulation across fields. Lacking process data describing contributions, evaluation subjects tend to use visible indicators like publication and project completion numbers within the same evaluation cycle. This encourages opportunistic behavior, with some researchers following trends for promotion and rewards while diligent, problem-tackling frontline talents go unrewarded, severely undermining confidence and enthusiasm. Fewer talents are willing to study "bottleneck" problems, hindering innovation vitality. Efforts should be made to achieve precise disciplinary affiliation, correct categorical and tiered evaluation, and authentic, sufficient evaluation data to strengthen responsible evaluation.

3.5 Problems with Evaluation Environment

Talent evaluation involves environmental factors including science and technology, economy, and culture that constrain and influence each other. Rapid S&T development requires evaluation subjects to correctly grasp trends and frontiers, significantly affecting results. Economic development levels determine application value of results, substantially influencing outcomes. Social culture, including relationship culture and contract culture, deeply affects evaluation. While peer review is widely accepted, implementation sometimes deviates, causing public skepticism due to relationship-based, acquaintance-based, or convenience-based evaluations that depart from evaluation essence. Talent evaluation should objectively and fairly identify positions and development potential within the talent ecological community based on capability, quality, contribution, and value. However, the aforementioned reasons cause evaluation essence to shift, inevitably producing inaccurate results. Maintaining environmental factors' non-influence on evaluation subjects and objects to achieve responsible evaluation is a crucial problem requiring resolution.

3.6 Problems with Evaluation Results

Evaluation results are important bases for implementing talent incentive policies. When implementing reward and punishment regulations, organizations should consider not only whether researchers have completed assigned workloads but also whether rewards match research contributions. If organizations only evaluate professional levels based on workload and social influence without considering practical significance, it encourages undesirable practices like "seniority-based promotion" and "title-only theory." Correctly understanding and utilizing evaluation results demonstrates respect for knowledge creation and talent labor, helps organizations express appreciation for S&T talents, and positively affects evaluation feedback. This enhances talents' self-identification and national confidence in China' s research environment, motivating diligent study while reminding both evaluation subjects and objects of their roles and responsibilities, providing solid review materials, establishing fair standards, ensuring authentic and effective results where all participants share responsibility, truly achieving responsible evaluation.

4 New Trends in S&T Talent Evaluation

New concepts and technical methods provide new perspectives for S&T talent evaluation. Supported by big data, artificial intelligence, blockchain, and other advanced technologies, S&T talent evaluation can achieve whole-process, full-cycle, all-dimensional, all-element, full-dynamic, and multi-category evaluation, becoming more scientific and rational.

4.1 Data-Driven Evaluation Mechanisms

S&T talent evaluation should avoid single-dimension generalization; different fields require different standards. Traditional systems rely primarily on scientometric quantitative methods and peer review qualitative methods. Research is a continuous process where talents grow. Evaluation materials typically include papers, patents, publications, and autobiographies, but textual information alone cannot verify process authenticity. To provide more supporting evidence and ensure scientific evaluation, knowledge network-based methods have emerged [74].

Research activities are multi-threaded, non-linear, and dynamic, requiring extensive references. Knowledge networks can record associations between reference information and research results, incorporating research process information and credibility into evaluation databases to provide decision-making references and improve precision. With the big data era, combining traditional methods with AI has become inevitable. Peng Zhen [75] proposed a quality-oriented evaluation method based on big data, transforming talent discovery from passive to active. Wang Haiyan [76] suggested establishing "S&T ID cards" for talents using blockchain technology, where research teams and academic systems can supplement materials, providing more detailed records. Data-driven meth-

ods feature large data volumes and fast processing speeds, generating user logs that transform scattered, irregular data into ordered, structured data, providing commercial and social value for evaluation decisions.

4.2 Open and Shared Evaluation Methods

In the context of open resource sharing, peer review occupies an increasingly important position. However, traditional peer review has drawbacks including weak expert constraints, industry interest bundling, and potential bias. To avoid these deficiencies, academia has adopted open peer review, non-selective review, and preprint models [77].

Beyond peer review, with social media popularization and knowledge sharing trends, scholars have expanded academic discussion to online platforms. Researchers post research updates and evaluations of hot topics on social media, whose fast feedback, wide dissemination, transparency, and directness provide new perspectives for achievement evaluation. For example, Funk proposed a disruptive index based on network models to measure research novelty [78]; Wu used the disruptive index to study research outputs and preferences across different team sizes [79]; Liu Xiaohui introduced the disruptive index into achievement evaluation [80]. Information technology provides broader perspectives and more complete research process data to enhance result usability and authenticity, while open sharing offers ideas for timely understanding of domestic and international frontiers and stakeholder evaluations, ensuring comprehensive and three-dimensional individual talent evaluation.

4.3 Representative Achievements Evaluation Mechanism

In early evaluation system development, representative achievements served as important references for academic leader selection, key laboratory applications, and project management. The 2003 revised *National Key Laboratory Assessment Rules* specified representative achievement forms including academic papers, monographs, key innovative technologies, and authoritative scientific data for decision-making. It was recognized that research achievements vary significantly in quantity, research paths, and quality. Evaluating all achievements would consume substantial resources and unconsciously reinforce the stereotype that "more quantity equals greater contribution." Representative achievements best reflect research performance and levels, 被视为 the most prominent and influential accomplishments of individuals or teams. Using their "equivalence" can assess innovation capability, shifting from quantity-dominated to quality-dominated evaluation, truly achieving the effect of "evaluation promoting prosperity."

Under policy guidance, S&T talent evaluation systems seek breakthroughs to reverse quantity-oriented patterns and combat "quick success" mentality. In 2010, Fudan University first introduced representative achievement evaluation to faculty appointment and assessment [81]. Subsequently, this mechanism gradually

entered public view, with universities exploring and practicing it. The 2017 *Ministry of Education Talent Development Guiding Opinions* explicitly stated the need to establish and improve representative achievement evaluation mechanisms, laying the foundation for widespread application.

4.4 Responsible Evaluation Mechanism

To balance fairness, transparency, and accuracy with disciplinary differences, indicator diversity, and evaluator subjectivity, evaluation systems require improvement, correction, and supplementation. Against the backdrop of flourishing innovation, the concept of "responsibility" has been introduced into research innovation and services to meet synchronized management development requirements and reflect people-oriented thinking. "Responsible"implies precision, standardization, effectiveness, and diversity, aiming to view science and technology from public interest and social ethics perspectives, giving rise to concepts like "responsible metrics," "responsible innovation," and "responsible evaluation."

In 2015, Professor Wilsdon first proposed "responsible metrics" [82], reminding academia to re-examine the relationship between quantitative indicators and research evaluation. Research practice is part of social activity; evaluation results themselves incorporate social shaping of achievements [83]. Overemphasizing quantitative indicators, such as whether papers are published in core journals or citation frequencies, can shift the academic atmosphere from long-term focused research to short-term "fast-food" results, potentially causing misevaluation of promising, practice-oriented achievements.

As awareness grows that evaluation should shift from quantity and journal prestige to quality and social contribution, responsible evaluation mechanisms outlined in the *San Francisco Declaration on Research Assessment*, *Leiden Manifesto*, and *The Metric Tide* report [84] have gained increasing endorsement.

In February 2018, the *Guiding Opinions on Deepening the Reform of Talent Evaluation Mechanisms by Category* explicitly stated "adhere to evaluating talent based on ability, actual performance, and contribution, overcoming tendencies to focus only on academic credentials, seniority, and papers." In June 2018, President Xi Jinping emphasized the importance of S&T talent development systems and breaking stereotypes of evaluating heroes based on papers, titles, and credentials. Later that year, the Ministry of Education issued the *Notice on Breaking the "Four Onlys"* and *Notice on Breaking the "Five Onlys,"* aiming to break the "quantitative dominance" pattern and provide fertile ground for innovation.

Beyond seeking changes in quantitative indicators, qualitative methods continuously optimize. Peer review has always been crucial for S&T talent evaluation. First, research projects encounter many unpredictable factors whose impact and appropriate responses cannot be determined by quantitative indicators alone, requiring domain experts' judgment based on knowledge and experience—something pure quantitative evaluation cannot achieve. Second, research pursues

innovation, and even projects in the same discipline have different innovation points that quantitative indicators may not fully capture, ensuring consistent sensitivity to different projects' innovations. Therefore, expert review mechanisms represented by peer review are indispensable and must continuously deepen reform to meet research management needs.

Based on the guiding ideology of "comprehensively deepening reform,"in 2021 the Management Science Department innovatively introduced the "Responsibility, Credibility, Contribution (RCC)" review mechanism into special evaluations [85]. The RCC mechanism imposes higher requirements on both subjects and objects, addressing both evaluation and feedback processes. It emphasizes matching reviewers' professional scope with evaluation objects, requiring thorough understanding and detailed opinions to ensure credibility and authority. A "credibility file"for review experts is established based on participation enthusiasm and result effectiveness, encouraging more responsible, fair, and professional participation. Research shows that introducing RCC in categorical evaluation positively impacts timeliness, quality, feedback, and reviewer attitudes [86]. Although pilot tests demonstrated RCC' s feasibility and advantages, issues emerged including overly positive comments with fewer constructive suggestions [87] and incomplete reviewer understanding of the RCC mechanism [88]. In 2022, the RCC mechanism will expand pilot scope, optimizing itself in practice, increasing visibility, and encouraging more experts to participate in innovative evaluation mechanisms for more objective and responsible decision-making.

Responsible evaluation embodies modern reflection on the relationship between technological progress and human society. Any innovation faces risks and challenges. In perfecting S&T talent evaluation, subjects should maintain alignment between innovation direction and social needs, matching actual contributions with rewards and punishments. Only by keeping the ship of people' s interests and technological innovation heading in the same direction will China' s technological giant ship sail farther and more steadily.

5 Conclusions and Recommendations

S&T talents are builders driving national progress and high-quality transformation, representing important human resources competed for globally. As innovation gains increasing attention and talent cultivation scales expand—evidenced by large talent numbers, broad disciplinary coverage, and refined research directions—S&T talent evaluation has become a crucial basis for talent selection and resource allocation, making scientific evaluation an important issue requiring attention.

This paper uses web surveys and content analysis to examine literature and policy texts on S&T talent evaluation from 2010-2021, exploring new landscapes and trends through policy descriptions, theoretical foundations, evaluation orientation shifts, and methodological changes. Current evaluation systems still have problems including coarse-grained talent classification, incomplete evalua-

tion dimensions, and suboptimal evaluation elements and processes. To address these issues and maintain evaluation integrity, new trends are emerging rooted in big data, powered by open sharing, and signaled by representative achievements and responsible evaluation.

Through reviewing existing policies and literature, future work will focus on implementing detailed evaluation criteria, measuring social impact of achievements, and practicing post-evaluation to further improve the S&T talent evaluation system.

Author Contributions: FENG Guohe was responsible for overall conceptualization and revision; CHEN Chenyan drafted the initial manuscript.

Note: Figure translations are in progress. See original paper for figures.