European Forum for Studies of Policies for Research and Innovation

CSIC

Instituto de Políticas y Bienes Públicos

Early Career Researcher Conference (ECRC)

Science dynamics and research systems:
The role of research for meeting societal challenges"

Madrid, 8-9 April 2013

The "Balancing" Role Played by the Evaluation Mechanisms of Key National Extra-university Research Institutes in their Own Strategic Development and in Support for National Research Policy-making ---The Cases of German Max Planck Society (MPG), the Helmholtz Association (HGF) and the Chinese Academy of Sciences (CAS)

Junwen Luo (MSc),

STePS, University of Twente

The "Balancing" Role Played by the Evaluation Mechanisms of Key National Extra-university Research Institutes in their Own Strategic Development and in Support for National Research Policy-making

---The Cases of German Max Planck Society (MPG), the Helmholtz Association (HGF) and the Chinese Academy of Sciences (CAS)

Junwen Luo (MSc), STePS, University of Twente

Abstract: Extra-university research institutes, as relevant academic contributors to the process of maintaining a national innovation system, are required to guarantee scientific quality for their own strategic development and provide guiding evidence for national research policy-making. With this dual mission, one could expect their evaluation mechanisms to contribute to a "balancing" function among researchers and various stakeholders in cases of diverging interests, disagreements or debates. Here, evaluation information might help to communicate, negotiate and mitigate potential struggles, tensions and controversies, like funding priorities and thematic research priorities. The design of the evaluation system with considerations of both the institutes themselves and key stakeholders will be explored empirically through the study of the cases of the German Max Planck Society (MPG), the Helmholtz Association (HGF) and the Chinese Academy of Sciences (CAS). A comparative analysis will give insights into both existing problems and future prospects, by reviewing the development dynamics of the evaluation experiences in the past fifteen years and analyzing the current situations.

Key words: Balancing Evaluation mechanism; research institutes; stakeholders; strategic development; research policy-making

Introduction

Evaluation procedures are utilized as a way of measuring the scientific and technological quality or the potential socio-economic impacts of publicly funded research (Becher and Kuhlmann, 1995; Bozeman and Melkers, 1993, Georghiou, 1995; Rip, 1990). However, there are numerous conflicting goals that are considered important by different agencies and stakeholders. There is no common ground for enabling 'key policymakers, managers, and staff' to come to an agreement about which goals should have priority (Palumbo and Hallett, 1993). In view of this it has been suggested to use evaluation as "a medium for the moderation of struggles, controversies and negotiations in the Science and Technology policy area" (Kuhlmann, 1998). Attention to stakeholders has gained prominence for both practical and ethical reasons, and involvement of key stakeholders is presumed to enhance the design and implementation of evaluations and the use of evaluation results in decision-

making (Bryson & Patton, 2010). No evaluation can answer all potential questions equally well. Assuming that innovation processes are uncertain, complex, linked to social and technical imponderabilities, and likely taking a different course from what is planned, the ideal way is to design the evaluation mechanism with considerations of both a research institute itself and other main stakeholders, depending on their respective importance and related evaluative motivations and relevancies.

Research evaluation takes place in a situation where autonomous but interdependent actors (research institutes, government, innovation-related ministries, science community, research funding councils, other funding agencies, societal stakeholders, and industries) are involved in struggles, tensions and controversies needing to be "balanced", like funding priorities and thematic research priorities. Involvement of key stakeholders is to enhance the implementation of evaluations and then the use of evaluation results in decision-making. Therefore, an evaluation process with considerations of both the research institute itself and other main stakeholders can play a balancing role by communicating, negotiating and mitigating potential struggles, tensions and controversies among various stakeholders. The balanced result could reflect wider and deeper concerns from various perspectives and contribute to institute's strategic development and providing guidance for national research policy-making. This kind of dynamics in the long term would facilitate reflections of actors and influence the next-round evaluations positively.

The empirical references will come from the three cases: the Chinese Academy of Sciences (CAS), the Max Planck Society (MPG) and the Helmholtz Association (HGF). The German research system embraces a high level of segmentation (H. Thomas & S. Kulmann, 2008) with large extra-university sector including institutes of the Helmholtz Research Centers (HGF) and the Max-Planck Society (MPG). Interestingly, the Chinese Academy of Sciences was founded by the example of Russian National Academy of Sciences which draw on German practices. To some degree , The CAS is more like the combination of the two German organizations and certainly has the motivation to learn their evaluation experiences.

The Max Planck Society (MPG) and the Helmholtz Association of German Research Centers (HGF) are two of the four main non-university research institutions (the other two: Gottfried Wilhelm Leibniz Science Association WGL, and the Fraunhofer-Gesellschaft FhG). The MPG institutes are well positioned in the field of outstanding basic research. It is convinced that only scientists of equal standing-peers are in a position to competently assess their colleagues' work, so the world's best scientists with professional competence, high reputation and wide-ranging experiences are sought to become its Scientific Advisory Board members for evaluation purpose. The rotating system of Board member replacement also strikes a balance between precise knowledge on the one hand and excessively closes ties between Board and institute on the other. Facilities of the Helmholtz Association (HGF) have the task of bridging the gap between basic research and applied research, whose original mission is directed at research on systems based on large-scale facilities and comprehensive

scientific infrastructure. As the only national and most influential research institutes system of China, the Chinese Academy of Sciences (CAS) has been working for national strategic development as a leading strength of governance of science and technology since it was founded in 1949. Now the CAS focuses on the fundamental research itself on the one hand (similar to MPG), and cooperates with ministries, research councils to make national S&T research policies orientated by strategic development of the whole country (similar to HGF). So the overall function of the CAS is like a combination of MPG and HGF, which means its operation and evaluation mechanism improvement could refer to both of the MPG and the HGF developmental dynamics.

Definitions

Evaluation mechanism here refers to the surroundings and processes of the evaluation activities from inputs (offer and demand analysis, organization work like selection and coordination of evaluators, contract procedures, designing evaluation indicator system) to outputs (reports, evaluators' recommendations, stakeholders' feedbacks). The "balancing" role is to consider the interests and potential controversies of internal and external stakeholders into the designing of evaluation mechanism, to communicate, negotiate and mitigate the potential controversies and finally positively influence the research institutes' own strategic development and providing references for national research policy-making.

Research question and sub-questions

My main key question is: how does this "balancing" evaluation mechanism look like in theory and how can it be identified and assessed in practices? Key sub-questions are as following:

1. The comparative analysis will concern the development dynamics of the evaluation mechanisms of the three cases in the past fifteen years. What are the most significant institutional changes concerning evaluation mechanism? Among the most important reasons to the changes, how much impacts are from the two aspects: the institutional strategic development and supporting the national research policy? Will the evaluation development also impact the two aspects in return?

2. What are basic and typical potential struggles and controversies among both internal and external stakeholders in evaluation practices? How are they considered in the designing of the evaluation indicators system of research institutes?

Stakeholders	
Stake or interest in the research institute: What do they want to get out of the institute, and/or What do they want the institute to produce?	Stake or interest in the research evaluation: What do they want to get out of the evaluation, and/or What do they want the evaluation to produce?

Fig 1. Basic stakeholder analysis. Source: Adapted from Bryson (2004a, 2004b)

	Estimate of Various Stakeholders' Initial Inclination Toward the Evaluation Recommendation		
How High Are the Stakes for the Stakeholders	Favorable	Neutral or Unknown	Antagonistic
High			
Moderate			
Low			

Fig 2. Mapping stakeholders' stakes and inclinations toward the evaluation's recommendations (adapted from Bryson and Patton (2001)

3. The 'balancing' evaluation mechanism will analyze and prioritize the stakeholders based on their different interests, powers, capabilities etc. before the evaluation procedures. We will see whether the information during the evaluative process or the afterwards results influence their original positions. If it does, can we infer that this 'balancing' role is working?

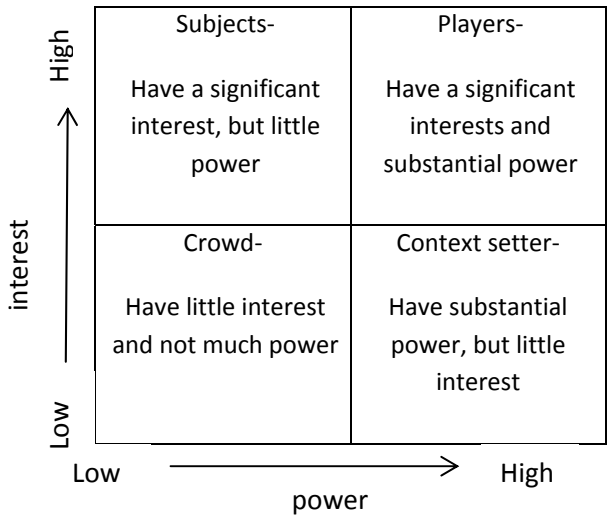


Fig 3. Power versus interest grid. Source: Eden and Ackermann (1998, p. 122)

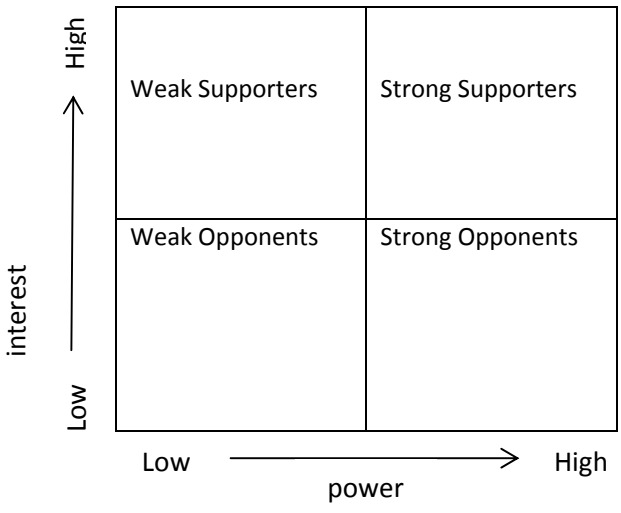


Fig 4. Evaluation recommendation support versus opposition grid. Source: Crosby, Brysom, and Anderson (2003)

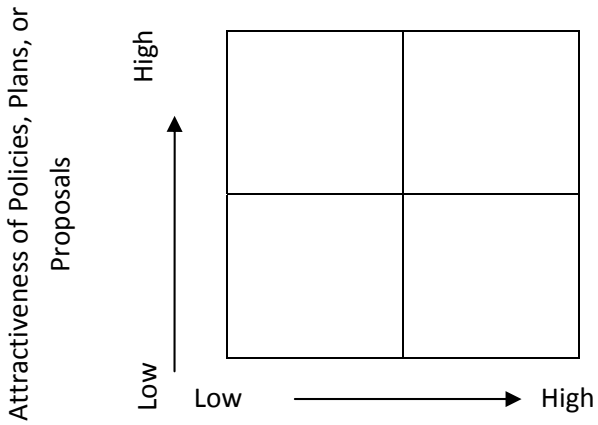


Fig 5. Recommendation attractiveness versus stakeholder capacity grid. Source: Bryson et al. (1986)

Stakeholders	Stake or Interest	Resources	Avenues of Influence	Probability of Participation and Manager of Doing so	Influence-as a Product of Resources and Participation	Implications for Implementation Strategy	Action Plan Elements
Supportive Stakeholders							
Opposing Stakeholders							

Fig 6. Recommendation implementation strategy development grid. Source: Adapted from Meltsner (1972), Coplin and O'Leary (1976), Kaufman (1986), and Christensen (1993)

4. what are constitutive elements of this "balancing" evaluation mechanism? Does the research institution would like to collect and reflect the concerns from all or part of internal researchers and external key stakeholders, for what purposes?

5. For the institutes' strategic development, does such "balancing" evaluation mechanism go for or against the scientific quality, like the freedom and creativity of scientists? For the reflection to national research policy-making, what kind of evaluative information should be selected and transferred efficiently and in which ways?

6. What are the differences of evaluation mechanism between Chinese and German cases? In which way and how the comparison of the cases would become insightful for all sides? Will German research institutes evaluation practices serve Chinese research institutions as a

blueprint for catch-up, or can they inspire the Chinese institutes the development of new tailored approaches in terms of evaluation mechanisms?

Hypotheses (expected results)

① During the past fifteen years, there are important institutional changes of the three organizations in the way that evaluation procedures are organized like: the restructuring reform of the HGF in 2001 and the related evaluation between 2002 and 2004; The system evaluation of the MPG in 1999; and the Knowledge and Innovation Program of the CAS since 2007.

② About the function of evaluation mechanism for the research quality assurance as a part of the institutional strategic development, the three organizations all agree. But about their roles in supporting national research policy-making, they probably take different attitudes and contribute to different degrees.

③ Although with the significantly different evaluation mechanisms, the three research institutes all experienced potential disagreements or debates in evaluative practices. Here is an estimation of the weighed concerns of internal and external stakeholders within these three organizations, generally reflecting how strong the stakeholders' interests in evaluative information are, how seriously their voices would be considered and taken into evaluative mechanism design.

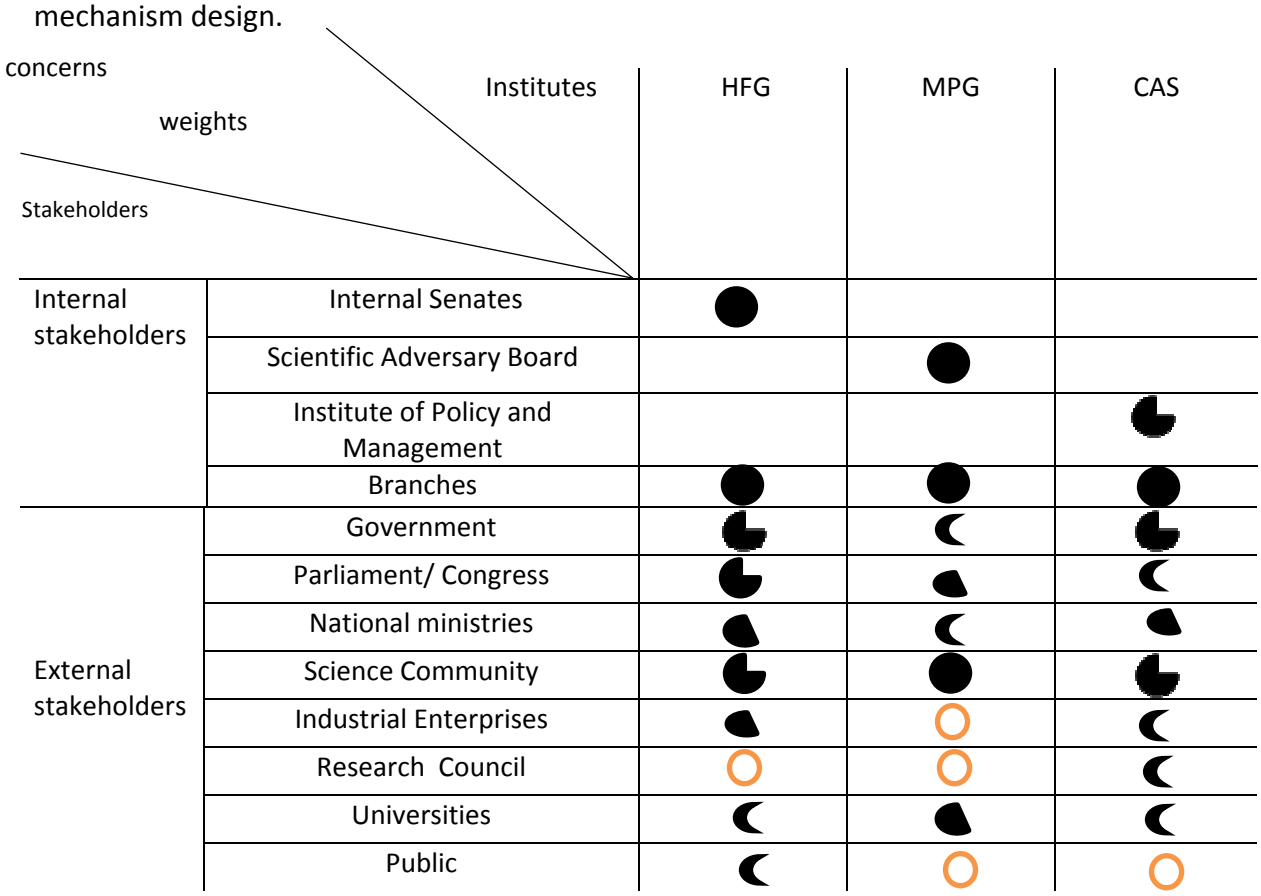
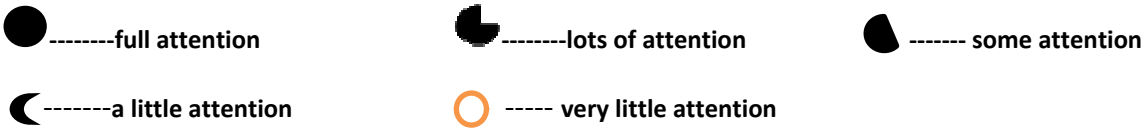


Fig 7. Estimation of different weighted concerns towards different stakeholders



④ The practices of the three cases prove that there are potential “balancing” functions of evaluation mechanisms through communication, negotiation, and even the solution of potential controversies of internal and external stakeholders even though they do not call them this way. The involvement of key stakeholders is to enhance the design and implementation of evaluations and then the use of evaluation results in decision-making.

⑤The CAS is still at an early stage of evaluation development but would like to improve it by learning German experiences. At the same time, with the more and more cooperation with the CAS, the two German research institutes also would like to share some widely-applicable experiences especially in the joint research programmes or cooperative organizations.

Case study backgrounds

The main empirical method is the case study of key national non-university research institutes in Germany and China: the HGF, MPG and CAS. Here are some basic information of the three national key extra-university research institutes.

items	Research institutes	employees	Years of development	Researchers under 45 years old (%)	Annual funding (billion)	Proportion of public funding (%)		Proportion of contract funding and others (%)
						Central government	Regional government	
HFG	18	32,698	44	?	3.4	70		30
						90	10	
MPG	80	13,000	60	?	1.4	84		16
						60	40	
CAS	94	43,000	63	75	?	50		50
						?	?	

Fig 8. Comparative general information of the three institutes

1 German cases: the MPG and HGF

MPG has mandate to expand the boundaries of knowledge pursuing basic research at the highest level. Peter Gruss, the president of the MPG, attributes the ability to conduct knowledge-oriented research in an open framework and without ties to specific applications to public funding and its high degree of trust. A sophisticated, gradually increasing effective and credible evaluation procedure has been established since 1970s, which now recruits external and critical scientists of high standing from all around the world. The globally visible success of the scientists who work at the MPG is based on the broad scientific freedom they enjoy as well as the sophisticated quality assurance procedures. Now its budget of 2% of the

total expenditure on research and development in the Federal Republic of Germany requires a manner that supports cutting-edge international achievements in basic research and strict quality criteria to meet international standards. MPG institutes are structured in departments, each of which is headed by a scientific Director, who is able to freely and independently select and carry out their research activities. The Society pays particular close attention (ex-ante) to its selection of Institute Directors and regular reviews (ex-ante) the results of their work in order to ascertain present potential. The ex-post evaluation is based on the research results for the past two years. The process of continuous evaluation is centered on the Scientific Advisory Boards (SAB), which were set up in the 1970s as permanent evaluation bodies at the MPG institutes, which makes a major contribution to discussions of future prospects and strategy. The MPG currently has around 750 SAB members, who are both critics and advisors to those whom they evaluate, drawn from leading universities and research institutions from all over the world. More than 300 experts participate in the up to 40 SAB inspection visits that take place each year. A well-founded evaluation depends on the sound knowledge of the scientific work in question followed with lots of detailed written reports. The status report, visit by the SAB, and report by the SAB are three essential steps in evaluation procedure. Another extended evaluation beyond individual institutes asks for comparison with one another and analysis in a national and international context, in which criteria applied are scientific achievement, the efficient deployment of resources and the institutes' medium-term future prospects. Besides, the Senate of the MPG, in addition to members, also includes representatives from politics and business. In this way, reports on the selective use of the Society's funds and the results of its quality assurance are presented simultaneously to the high decision-making body of the Society and to the funding providers. Bibliometric, citation or patent analysis have increasingly been used by experts and the general public as a tool for measuring quality and performance. But sometimes quantitative information and performance indicators can funnel valuable input into the process of quality assessment. If the evaluators are aware of the strengths and weaknesses of the indicators and are provided with a broader approach to policy relevant evaluation , they can even help to avoid costly errors in the assessment as well as in the subsequent decision making process. The evaluations conducted within the MPG make use of a wide variety of different data and indicators, e.g, budgetary plans, publication lists, citation and co-citation analysis, prizes awarded to the researchers concerned, etc, and make it one of the main obligations of the expert panels to achieve a consensus on the weight of the various indicators according to the special nature of the institutions subject to review.

As the largest scientific organization in Germany, the Helmholtz Association of German Research Centers (HGF) is a union of 18 scientific-technical and biological-medical research centers with about 70% of the budget raised from public funds. The official mission of the Association is "solving the grand challenges of science, society and industry". The HGF conducts research on highly complex systems using large-scale equipment and scientific infrastructures together with national and international partners, where scientists focus on

complex systems which affect human life and the environment. It contributes to shaping the future by linking research and technology development with innovative prospects of application and prevention, therefore identifies and works on major, pressing issues of society, science and industry by means of strategic and programmatic cutting-edge research. HGF has strong collaboration with the CAS over the past 30 years. Its Beijing office is responsible for information exchange and joint programs cooperation, like joint training PhD students.

In autumn 2001, there is a significant reform in the HGF. For the purpose and against the background of an increasingly close networking between knowledge-oriented basic research and industrial applications, the mission and work priorities of the Helmholtz were restructured and focused on six research areas (energy, Earth and environment, health, key technologies, the structure of matter as well as transport and space). All six research areas were evaluated between 2002 and 2004. About 350 internationally renowned scientists, more than half from abroad, participated in the evaluations. Experience with the evaluations and the ongoing programmes show that the new programme-oriented funding procedure proves to be valuable. All in all, the procedures have strengthened the responsibility of the Helmholtz Centres themselves and sharpened their profile as well as increasing national and international visibility. The core of the fundamental reform of the Helmholtz Association was the introduction of "programme-oriented funding". The objective is the development of a new quality of cooperation between the legally independent Helmholtz Centres. By means of the reform, the considerable potential and large resources of the Association are to be used more effectively and its orientation to performance and results is to be strengthened. Increased cooperation and competition between the centers are strategic objectives which will both be optimized. Within the framework of programme-oriented funding, the scientists of the Helmholtz Association develop programmes in each research area which combine contributions by research groups on central topics and different centers in an interdisciplinary way. The programmes are evaluated by international groups of experts and assessed for their scientific quality, competence and strategic importance. The results of the evaluation then form the basis for funding decisions taken by the Federal Government and the Länder. Overall, the HGF's structure is more similar to Chinese case, especially its strong orientation of national policy guidance for government. Or in other ways, the HGF faces more and wider stakeholders than MPG, so it might have wilder considerations from other voices than just the scientific community.

2. Chinese case: the CAS

The CAS has six academic divisions (Mathematics and Physics, Chemistry, Life Sciences and Medical Sciences, Earth Sciences, Information Technological Sciences, and Technological Sciences), over one hundred institutes, one university and one graduate school, located in 20 provinces and municipalities throughout China. The CAS has invested in or created over 430 science- and technology-based enterprises in eleven industries including eight companies

listed on stock exchanges. For such a huge research system, the operational control of institutes and promotion of communication among institutes of similar research fields require a comprehensive and clear evaluation mechanism as guidance. Now there is an Evaluation Division in the Headquarter responsible for all the evaluation planning of the CAS system, and an independent Evaluation Research Center carrying out evaluation activities under those planning, of which many evaluators or working staff are also the researchers in the Institute of Policy and Management (IPM) of the CAS. The dual identity may leads to better understanding and integration of evaluative theories and practices.

S&T evaluation in China is coming into the spotlight of the governmental functionaries, S&T administrators, academia and mass media. In the nearly 20 years of development of the institution evaluation, it has experienced from output evaluation (1993-1998), performance evaluation on academic achievements and level of accomplishment (1998-2004), to quality-oriented comprehensive evaluation focusing on S&T creativity and concerned performance in integrity (since 2005). In 1998 the Knowledge Innovation Program (KIP) was launched by CAS as a milestone. In recognition of the national strategic requirements and the world trend in science and technology, the process of conducting the KIP gave rise to the most profound and extensive adjustments in its disciplinary development and organizational structure. Scientific structuring and management innovation has led to the preliminary formation of a structure and mechanism that are both geared to a socialist market economy and aligned with international standards. Then in 2007, seven pilot institutes were selected to investigate modern R&D institutions system, and evaluation becomed one of the four thematic considerations (the other three: resource allocation, human resource, organizational model of R&D). There is empirical evidence (Daqun Zhang, Rajiv D. Baker, Xiaoxuan Li etc. 2011) shown that the productivity of the research institutes of the CAS has been improved by 12.5% after its implementation of the KIP, further decomposed into 8.8% attributed to technological progress and 3.3% to efficiency improvement. Results of regional analysis show that institutes in Beijing and Shanghai performing better than institutes in other cities during the same period.

There is a three-hierarchy evaluation model in this pilot study: indicators reflecting common characters at the bottom like yearly quantitative monitoring and yearly communication review; in the middle key indicators and benchmarks reflecting the characters of certain category of institute, which can be used in both diagnosing evaluation for individual institute and comparison between institutes of the same category; at the top qualitative evaluation by experts review with two stages of diagnosing evaluation for individual institutes and comparing evaluation among the institutes of the same category. The selection of key indicators: based on the orientation and characteristics of institutes: six key indicators for each pilot institute, three common key indicators and eight specific key indicators. Four methods to set the benchmark: 1) the more the better; 2) different standards given based on international comparison; 3) benchmark set according to the development level of institutes; 4) judged by experts experience.

indicators	The method to set benchmark
S&T talents	Make use of the definition of quantitative monitoring on innovation capacity about the S&T talents
Per capital funding	With reference international research institutes and the real per capital budget of this year
Awards	More the better

Fig 9. The three-hierarchy evaluation model

Through the study of seven pilot institutes, key indicators and benchmarks regarding different types of institutes were proposed, including key quantitative indicators and benchmarks as well as key qualitative indicators and anchoring method. Thereby, three hierarchy evaluation model has been set up. The new model puts more emphases on the character of evaluation by category and individual features of institutes. In summary, this model has developed Quality-orientated Comprehensive Evaluation system in methodology.

Common key indicators	S&T Talents
	Per capital funding
	Awards
Specific key indicators	Invited reports of important conference or workshop
	Accomplishment of Major Tasks
	Construction and operation of large research facilities
	Platform of resource and data
	Intellectual property
	High quality papers
	Engineering application and demonstration
	Advisory reports

Fig 10. The three-hierarchy evaluation indicators

As one of the main three S&T Evaluation System of the CAS (the other two are self-evaluation for supporting strategic management and specific evaluation for demonstrating policy orientation), institute evaluation includes the overall quality evaluation, innovation capacity monitoring and input-output analysis of institute. The quantitative monitoring on innovation capacity, a comprehensive indicators system referring to five aspects with 24 fundamental indicators, adopts both the lateral comparison among institutes for providing status information to headquarter and the longitudinal comparisons for self-monitoring. The annual self-evaluation form of the institutes with 31 indicators is not public. Neither is the report conclusion as guidance for public money allocation. The qualitative evaluative methods mainly consider peer review on quality and capability, as well as review of advisory board, expert judgment about the strategic direction, resources allocation, efficiency and performance management, which vary according to the different scales of the units.

items	subdivided	Evaluation grading
SelfSelf--evaluation : Supporting strategic	Situation Analysis of CAS	
	Assessment of S&T layout	

management	Plan and policy evaluation	
	Evaluation of S&T Innovation Base	
Institute Evaluation: to promote competition & development	Overall quality evaluation	
	Innovation capacity Monitoring	
	Input-output analysis of institutes	
Specific Evaluation: demonstrating policy orientation	Project evaluation	
	Evaluation of human resource	
	Evaluation of Laboratory	
	Assessment of policy guidance	

Fig 11. An institute self-evaluation of CAS (exchanging with UK experts 2008)

Score	International cooperation projects	Post holding in international academy	International academic exchange	Capacity of attracting S&T resources internationally
Institute A				
Institute B				
Institute C				
Institute D				

Fig 12. "Traffic light" grading system of innovative capacity evaluation

There is no single, correct construction of any program's goals, implementation, or impact. But China's current management evaluative system is too simple and administrative (Dongbi, 2011). The first international evaluation of the CAS took place in its Physics Theory Institute in 2005, of which the president of the evaluation committee is American physicist and Nobel Prize winner David Gross. One of the committee members Chengtong Qiu suggested that the existing unclear and unfair evaluation mechanism would probably mislead the young researchers to take personal networking rather than hardworking as the qualification for successful research.

Ongoing Interviews of field work

The comparative analysis are being investigated by field work. Now I am at the stage of the first round of interviews for general understanding of the existing evaluation mechanism of the three organizations. The next round is expected to find out the specific comparative subjects, like the evaluation of Physics Institutes with similar orientation of the very basic research, to explore the potential "balancing" evaluation mechanism.

Here is the first-round interview guidelines table. I just finished the CAS part last week and am sorting out the interview data, some of which have not been referred to in this draft but will be discussed in the coming presentation of ECRC in Madrid.

	Interview questions to the three organizations
Hypothesis ? Development Dynamics	Brief introduction of the existing evaluation mechanisms of the research institutes
	During the past fifteen years, What are important changes in the way that

	evaluation procedures are organized? For example: the restructuring reform of the HGF in 2001 and the related evaluation between 2002 and 2004; The system evaluation of the MPG in 1999; and the Knowledge and Innovation Program of the CAS since 2007
Hypothesis □ Two Missions	Does evaluation positively impact institutional strategic development as the research quality assurance?
	Is the evaluation taken seriously by ministries, like as a citation for design of national research policies and programmes?
Hypothesis □ Stakeholders Consideration	Who are the general internal and external stakeholders in evaluation practices?
	Do you classify and prioritize stakeholders according to their interests, powers and capabilities? An example would be appreciated.
	In the given example, are there potential disagreements or debates from various stakeholders?
	How do the debates or potential tensions from stakeholders influence the evaluation procedures or results?
Hypothesis □ The Role of Evaluation mechanism	Have you considered the involvement of the stakeholders in evaluation mechanism design of the next round evaluation?
	Do you agree that their involvement help the communication and negotiation?
	Do you agree that a more comprehensive evaluation mechanism considering stakeholders' interests is beneficial to the long-term development of your institution?
	Do you think such comprehensive evaluation result will be useful for the ministries to design national research policies or programmes?
Hypothesis ⑤ Experiences Sharing and Learning	For the two German institutes, would you like to share some widely-applicable experiences especially in the joint research programmes or cooperative organizations?
	For the CAS, will the German research institutes evaluation practices serve Chinese research institutes as a blueprint for catch-up, or can they inspire the Chinese institutes the development of new tailored approaches in terms of evaluation mechanism?

Fig 12. First round interview guidelines

References

Stefan Kuhlmann, 1998, Moderation of Policy-Making? Science and Technology Policy Evaluation Beyond Impact Measurement-The Case of Germany, Evaluation, 4(2), 130-148

Thomas Heinze, Stefan Kuhlmann, 2008, Across Institutional Boundaries? Research Collaboration in German Public Sector Nanoscience, Research Policy, 37, 888-899

Johb M. Bryson, Michael Quinn Patton, Ruth A. Bowman: Working with evaluation stakeholders: A rational, step-wise approach and toolkit, Evaluation and Program Planning, 34(2001)1-12

Korthonen P, Tainio R, Wallenius J, 2001, Value efficiency analysis of academic research. European Journal of Operational Research 130, 121-132

Suttmeier RP, Cao C, Simon DF, 2006. Knowledge innovation and Chinese Academy of Science. Science 312, 58-59.

Daqun Zhang, Rajiv D Banker, Xiaoxuan Li, Wenbin Liu, 2011, Performance impact of research policy at the Chinese Academy of Sciences. Research Policy 40, 875-885.

W. Krull, 1995, The Max Planck Experience of Evaluation, Scientometrics, 34(3), 441-450.

B. Koester, 2002, Example: Helmholtz association of national research centres, Germany

Research, Innovation and Technological Performance in Germany Report 2012.

Evaluation: the procedures of the Max Planck Society from its website.

Xianxuan Li, Bing Shi, Jianzhong Zhou, 2009, A study of the model and methodology for institute evaluation in CAS, Evaluation Conference in Orlando, Florida.

Arie Rip and Barend van der Meulen, 'Science policies as principal-agent games: institutionalisation and path-dependency in the relation between government and science', Research Policy, 27, 1998.

Dietmar Braun and David H Guston, Principal-agent theory and research policy: an introduction', Science and Public Policy, 30 (5), 2003.

Elizabeth Shove, 'Principals, agents and research programmes', Science and Public Policy, 30 (5), 2003, 371-382

Dietmar Braun, 'Who governs intermediary agencies" Principal-agent relationships in research policymaking', Journal of Public Policy, 27 (8), 1993

Dietmar Braun, 'Lasting tensions in research policy-making – a delegation problem', Science and Public Policy, 30 (5), 2003, 309-322

Elizabeth Shove, 'Principals, agents and research programmes', Science and Public Policy 30 (5), 2003, 371-381

Barend van der Meulen, 'New roles and strategies of a research council: intermediation of the principal-agent relationship', Science and Public Policy, 30 (5), 2003,323-336

Fetterman, D. M., S. J. Kaftarian and A. Wandersman (eds.) (1996): Empowerment Evaluation. Knowledge and Tools for Self-Assessment & Accountability. Thousand Oaks, CA (Sage).

Hwang, Byung Yong; Soon Cheon Byeon; Byung Tae Yoo: The Influence of the Government-funded Research Institutes (GRIs) Evaluation System on the Acceptability of Evaluation Results and Performance of the GRIs, Asian Research Policy, Volume 1, Issue 2, 2010 (September 30), 120-136.