

Aggregated Governance by R&D Evaluation Mechanism - Case Study of Chinese Academy of Sciences

Junwen Luo^{1,*}, Gonzalo Ordóñez-Matamoras^{1,2}, Stefan Kuhlmann¹

Abstract

This article aims at increasing understanding of the role of R&D (Research and Development) evaluation mechanism in mitigating conflicts of governance under the organizational framework of Publicly Funded Research Institutions (PRIs). There has been a series of governance reforms within the Chinese Academy of Sciences (CAS) in the past eighteen years. These vigorous reforms over the years have improved significantly CAS's stability of organizational research landscape and focused governance of complex R&D across varying institutes, disciplines and fields. Dynamics of evaluation mechanism, as a key instrument of governance reforms, has experienced four stages from zero, being purely quantitative towards strategic usefulness and categorized principles. Organizational governance of CAS has been learning to be increasingly aggregated from the institutionalization of evaluation mechanism and the growing stakeholders' communication and interactions. The roles of evaluation mechanism are for governing complexity and mitigating stakeholders' tensions in CAS, which perform varyingly at different levels of organizational framework. CAS study offers insights for governance of developing large-scale PRIs under similar organizational framework especially experiencing big reforms.

Keywords: organizational governance, conflicts, evaluation mechanism, learning, Chinese Academy of Sciences (CAS)

1. Introduction

Organizational actors as the main elements of an innovation system produce, adopt, diffuse and use knowledge and technologies (Borras & Edler, 2014). Development of scientific research is constantly shaping the organizational governance of PRIs under national and institutional contexts. Within national research landscape, PRIs have defined organizational missions from which all institutional arrangements derive involving complex R&D within member institutes under the organizational umbrella. Conflict of governance is originated from divergent concerns of stakeholders on three classified aspects: strategy, finance and operation. PRIs are increasingly required to provide reliable evidence for governance

effectiveness, for which evaluation now takes more and more responsibility. The evaluation mechanism as a whole involving all evaluation activities under the organizational framework is expected by the authors to consider, reflect and mitigate the governance conflicts.

Chinese science gets mass transformation with huge public investment, and the focus and challenges of R&D spending and evaluation have been addressed by Nature and Science journals (Cyranoski, 2014; Michelle, 2015; Sun & Cao, 2014). Particularly, the reforms of governance and evaluation within CAS, as a microcosm for the country, reflect the previous difficulties and new aspirations. To give an overview

¹ Science and Technology Policy Studies (STePS), University of Twente, Drienerlolaan 5, 7522 NB, Enschede, Netherlands;

² Facultad de Finanzas, Gobierno y Relaciones Internacionales-FIGRI, Universidad Externado de Colombia, Bogotá, Colombia

*Corresponding author: j.luo@utwente.nl

of the history of governance and evaluation, CAS established in 1949 has been the unique and dominated PRI in national S&T (Science and Technology) agendas and strategies. After the Open-Up Reform in 1978, China gradually developed out of the planned economy approach and transformed towards a market economy in the 1990's (Liu, Simon, Sun, & Cao, 2011). Accordingly, CAS has expanded significantly in scale, research quantity and quality, ranking as the sixth place among international research institutions in Nature Publishing Index 2013. Now it has over 60,000 employees from a centralized Headquarter with 6 academic divisions, 13 regional branches, 104 research institutes, 3 universities, over 400 invested enterprises in 11 industrious sectors including 8 listed companies, and other supporting units located throughout the country (website, 2015). "Academicians" as Individual Membership of the Academy represents the highest glorious title domestically for CAS scientists. The 104 member institutes under the umbrella conduct diverse research across various fields (like Health, Environment) and disciplines (like Physics, Chemistry). Meanwhile, CAS's governance became very complex with tensions from various actors. They were the fundamental reasons for the following remarkable reforms in the past two decades: the "Knowledge Innovation Programme" (KIP) from 1997 to 2010, the "CAS Innovation 2020" since 2006, The "Innovation 2050" since 2009, and the "Pioneer Action Plan" (PAP) since 2014. Overall, these reforms aim at simplifying the complexity and increasing the efficiency of CAS system strategically and financially. The evaluation, as a keyword of each reform, has been designed and implemented to address such complexity experiencing several milestones on those stages. Since 1990s the concept of evaluation entered into the spotlight of the CAS's R&D administrations (not even governance). In the previous period without evaluation, centralized

power of Headquarter and seniority of experienced professors jointly decide research direction and resources allocation. Evaluation of individual funding-oriented programmes and projects always take main responsibility of research quality. The governance reforms have driven the institutionalization of evaluation from zero, quantitative indicators dependence, introduction of international peer reviews, to the current Major Outcome orientations. The evaluation use has become far beyond justifying public investment and impacting resources allocation. What we explore are some functions realized in past practices and anticipated in the near future but unaddressed in literature on contextualization of evaluation usefulness. There are possibly some failures or side-effects in those dynamics which we did not study.

This article illuminates the role of evaluation throughout its contextual embeddedness and institutionalization dynamics in the CAS system. We firstly characterize the framework of multi-layered governance and highlight the embedded national research landscapes and institutional contexts. Then we develop exploratory typology of key stakeholders' tensions at various levels of decision-making. The functioning approaches linking evaluation and governance are through the institutionalized spread of evaluation procedures involving stakeholders' communication and interaction patterns. Eventually a series of roles are found working at the different layers of organizational governance to mitigate particular conflicts.

2. R&D Evaluation for Governance:

Theoretical Considerations

The conceptual discussion of this article is based on current debates on the role of evaluation as legitimacy instrument for effectiveness of governance (process and output). Beyond a way of

measuring R&D quality and socio-economic impacts of publicly funded research, policy evaluation procedures contributions to the moderation of policies in negotiation system which considers the heterogeneous struggles and controversies of the participating actors (Kuhlmann, 1999). PRIs are confronted with a multitude of always potentially controversial expectations (Dahler-Larsen, 2011) internally and externally: funding bodies, scientific communities, research funding councils, collaborating universities and enterprises, and general public. Evaluation thus has to be moved to the forefront of organizational thinking and behavior (Sanders, 2003). Various evaluation activities must be mainstreamed and integrated into a mechanism by organization for systemic arrangements.

How does this governance tool function? The focus to explain the justificatory use and impact of evaluation needs to be on the evaluating organization and its conditioning factors, rather than the evaluation itself (Hojlund, 2014). In a new mission-oriented policy context, the usefulness of evaluations as learning tools becomes relevant as the way current evaluation practices address broader impacts and issues (Amanatidou, Cunningham, Goek, & Garefi, 2014). Contextualization of evaluation design and usefulness were called for (Edler, Berger, Dinges, & Goek, 2012) from the practice of evaluation in innovation policy in Europe. Simon (2013) demonstrated that peer-review based evaluations can be an authoritative source of validation employed by scientific institutes to develop or optimize institutional strategies. Recommendations are about how institutional evaluations may be further developed and how processes of organizational transformation can be supported in the research landscape (Simon & Knie, 2013). If evaluative information and services are expected to become regular and sound materials for decision-making on strategic, financial and operational aspects, they must

be produced, transformed, negotiated, interpreted and reported (Kuhlmann & Edler, 1998) among participatory stakeholders with varying interests and powers.

Some tension in research institutions has been raised: the balance between exploitation and exploration of resources (March, 1991), and the balance between merit and patronage basis for research funding allocation (Nowotny, 2014). But we find roles of evaluation of PRIs in mitigating those tensions within national and institutional contexts are largely unaddressed while learning from institutionalization of evaluation mechanism is still underexploited. The thinking of evaluation is formed and shaped according to the rational, learning and institutional models (Dahler-Larsen, 2011) where most evaluators subscribe to the positive vision of the learning organization. Some balancing activities vis-a-vis stakeholders at different layers of organizational framework have responded to this development. The learning capacities of PRIs, such as adaptation in environmental innovation context taxonomy (Crow & Bozeman, 1998) and self-correction flexibly (Dahler-Larsen, 2011), could be reflected in the dynamics of organizational governance as well as the institutional spread of evaluation. Institutional learning through R&D evaluation mechanism is now being accumulated into mainstreamed channels of organizational governance.

3. Methodology

This article is derived from one case study of a larger PhD study¹ with a conceptual framework grounded in neo-institutionalism principles. The structure of this article, section by section, represents the logic and rationale of the conceptual framework. We design a three-level organizational framework based on literature, policy documents and also realities, involving central supervisory body at the macro-level (L1), institutional performance at the

meso-level (L2) and individual research at the micro-level (L3). Organizational governance is operationalized in three aspects: strategies, funding and operation with generalized dimensions like organizational missions, strategies, funding framework etc. Conflicts of governance are categorized in a three-by-three matrix according to the three aspects and key stakeholders at the three levels.

We reviewed the past eighteen years' dynamics of governance and evaluation of CAS through literature and the published policy documents. Based on that, qualitative data were collected and interpreted from in-depth and interactive individual interviews when we testified the theoretical concepts, and discussed their experiences and perceptions. We visited the Headquarter, Evaluation Research center and eight different research institutes of CAS with varying research characteristics, structures, and responses towards the reforms. Totally 28 in-depth individual interviews were taken (6 from L1; 14 from L2; 8 from L3) in 2013 at their work places, involving senior evaluation researchers and practitioners, Institute Directors, Heads of the Departments of institutes: Departments of Research Management, of Human Resources, of Education, of Technology Transfer, professors and early career researchers. And 6 external researchers on CAS's evaluations were consulted.

4. National and Institutional Contexts of Organizational Governance

Dimensions of organizational governance (mission, strategy, finance, operation, organizational change and learning) are all embedded in national and institutional contexts as the background of the following conflicts. We illustrate the governance reforms and evaluation dynamics separately in chronological order but would not relate them in one figure with a timeline. Structurally, several kinds

of working units at L2 with varying degrees of autonomy are governed by central Headquarter at L1: provincial branches, centers, legally independent (or dependent) institutes under the Director responsibility, and laboratories with large-scale facilities.

4.1 Complex Missions and Research Orientations under Simplification Processes

The governance has to do with the way in which politicized science works (Good, Vermeulen, Tiefenthaler, & Arnold, 2015). CAS has a mandate to comply with political tasks of the State Council, the highest government agency of China. In 2014 organizational missions of CAS were politically updated as taking the pioneering role in achieving Four First in China: "first in achieving leap-forward development of S&T, first in building high grounds for cultivating innovative talents, first in establishing high-quality S&T national think tank, and first in building internationally leading research institutes" (website, 2015). Its scientific missions aim at innovation capacity-building and sustainable development with emphasis on: 1) scientific frontiers for technology breakthroughs and industrial development; 2) scientific challenges to improve people's welfare while addressing resources and eco-environment; 3) high-tech challenges to improve international competitiveness and national safety (website, 2015). Besides, CAS builds up a complete educational system and serves over 70% of national S&T infrastructure and facilities. Overall, the complex missions and research orientations include basic research at frontiers; technological research, transfer and commercialization; nationally strategic research for long-term development; higher education; big-science facilities; scientific advice on S&T policy decision-making. Therefore, CAS gets the remark of "no organization in the world with so many functions", "being too big and unwieldy and lacking of inappropriate evaluation system" (Cao, 2014, 469).

The CAS's integrated R&D capacity with more synergies among institutes and fields were intended to improve by the governance reforms. Institutional contexts are constantly changing initiated by leaders at L1 which guide explicitly the adjustment of organizational structures and funding fields at L2, and scientific research at L3. KIP was the forerunner of categorizing and restructuring research institutes. 10 prioritized funding fields were created and the number of legally independent institutes was reduced to 99 from 123. There were 16 institutes reconstructed from the previous 46, 12 newly established and 7 renamed and 6 turning out as spin-offs (Report, 2012). About 25,000 researchers in CAS lost their tenured position ('iron rice bowl') to streamline the operation and to lower the costs of maintaining this massive community (Liu et al., 2011). These profound and extensive adjustments resulted in the preliminary formation of a structure and also aligned with international standards for focused resources allocation. And collaboration across the old fields and institutes was encouraged. Those scientific restructuring and management innovation led to the productivity of CAS institutes improved by 12.5% after the KIP, further decomposed into 8.8% attributed to technological progress and 3.3% to efficiency improvement (Zhang, Banker, Li, & Liu, 2011).

"Innovation 2020" emphasized indigenous innovation of institutes by another portfolio of stimulating policies based on KIP. Its 16 Major Programmes cultivated over 2000 outstanding S&T talents, invested costly but high-quality research facilities for international competitiveness and reputation, and transferred over 120,000 leading young talents from universities to industries. The "Nice Shifts" of "Innovation 2020" restructured the units guided by distinct categories of research orientations. And 17 new research fields were proposed to map within "Innovation 2050: China's S&T future".

A substantial reaction from Headquarter was its restructuring internal administrative Bureaus in 2013.

Previous Bureaus, originally as administration centers for individual institutes, had overlapping functions and lacked efficient coordination. New Bureaus of "Fundamental research", "Major R&D Programmes", "S&T Development" connecting with ministries, "Frontier Science and Education" linking to universities, "Strategic Planning" and "Finance and Support" dealing with internal affairs, were better classified by research orientations guiding efficiently relevant administrations. Additionally, the "Three-in-One" principle proposed at L1 has been penetrating down to L2 and L3 by strategically combining the three distinct roles of CAS: policy support, science exploration and education. They three coordinate with the new Bureaus to clearly position roles of each institute, to break down boundaries of units sharing similar roles and strengthen their collaboration (interview of Headquarter, 2013).

The latest PAP in 2014 CAS Congress was reported as "the biggest change in the Academy's history" with unprecedented structural reforms (Cyranoski, 2014, 468). The newly established modern innovative research units have been categorized into the four distinct research orientations according to their nature and goals of the current and future research activities (Poo & Wang, 2015). Each unit would be equipped with integrated resources and talents of the entire CAS. The first category is "Centers of Excellence" focusing on basic research in frontier sciences, and the areas where China has chances to dominate. These centers put together high level expertise selected from various institutes to solve great challenges, develop new research areas, and achieve breakthroughs with major impact worldwide. The second one "Innovation Academies" aims at larger scale of applied R&D outputs in technological and industrial areas with underdeveloped commercial potentials, each "Academy" in a particular field such as information technology. Creating such academy has to synergize the strengths of technology transfer of the existing institutes working in one field rather than just split those applicable units from the

institutes and then put them together in one academy (Poo & Wang, 2015). Big Science Facilities, as the third group, provide major large-scale S&T platforms open domestically and internationally. It links and makes collaborators around facilities easier for both fundamental and industrious research at different regions. The conceptually centralizing them means unified utilization, management and coordination rules considering geographic boundaries to improve the integral efficiency. The fourth one “Specialty Institutes” still in the design phase are devoted to researching specialized areas to assist local development, sustainability and people’s life, including preparation against natural disasters. Totally about 70 institutes are included in these four categories, each consisting of researchers across institutes. Some traditional institutes established by disciplines such as Institute of Mathematics would become departments of CAS universities.

4.2. Multi-channels Funding Programmes under Decreasing Processes

Some other missions are required from multiple programme and project funders. In 2012 CAS earned a gross income of about 37 billion RMB, with a ratio of 54:46 between block fund from governments (both the Ministry of Finance and regional governments) and competitive programmes and projects from various channels. Block fund is calculated according to official posts basically covering personnel staffing salaries and welfares. Competitive Programmes and projects are sorted to longitudinal (zong xiang) and crosswise ones (heng xiang) in Chinese language. Longitudinal ones are usually big-scale and disciplinary with strong policy orientations, from the long-term Programmes unpacked to subordinated projects funded by Ministry of Science and Technology (MOST) and National Science Foundation Council (NSFC) as majority and CAS Headquarter as minority. Crosswise projects are mostly funded by external enterprises, always by short-term contracts or entrepreneurial innovation platforms. In 2012 alone,

the whole Academy took 58 projects of “973 Program” (32% of the total), and 113 projects in “863 Program” (10% of the total), both funded by Ministries; and around 3,600 General Projects (12% of the total), and 125 Key Projects (21% of the total), both funded by NSFC (CAS, 2013). Not mention the tens of thousands small-scale projects funded by regional governments and industries. And the proportions of crosswise project funding vary among institutes but still remains small for the whole CAS. Our interviews show that multi-channels project funding was over 70% of total budget for every investigated institute.

There are other programmes beyond research funding. Attraction of outstanding talents from abroad has been taken by the “100 Talent Program” since 1994 with high-level targets and intensive supports. More than 1000 distinguished and young Chinese scientists from abroad have been successfully recruited as an important base for the development of CAS. The report on “The most Influential International Researcher” by Thomson Reuters in 2014 announced China ranked fourth with 111 researchers from Mainland among the total 3200 international researchers. However, low salaries for Chinese scientists are still considered as a key reason of being “not getting a good return on its investment, in terms of both basic-science breakthroughs and commercialization of those findings” (Cyranoski, 2014, 468). Since 1999 the Three Components Income has been applied: basic salary, position subsidy and performance award. But over 50% of the interviewed scientists admitted these three together could barely covered the basic living expenses. They had to apply for grants from multi-channel projects which accounted about 70% of their income. In fact, salaries of graduate students, retired personnel, and the costs of health care and other expenses all come from research funding. To secure sufficient funds, the researchers need to write many grant applications, leading to unnecessary competition, duplication, and fragmentation of the research programs (Poo & Wang, 2015). Besides,

under the pressure of anxious fighting for funds, scientists were away from collaborations because co-authorship is seen to dilute individual's achievement in the eyes of grant committees (Cyranoski, 2014).

To deal with these problems, Reform of Major Outcome Orientation has been focusing on decreasing the quantity but increase the quality of competitive projects. PAP has the task to supplement scientists' incomes by increasing the basic salaries and decreasing the proportion of research grants in the total income to 30% (Poo & Wang, 2015). The definitions of Major Outcomes are given differently by the six Academic Divisions across fields but closely following CAS's missions. The standards and requirements are specified by examples in each Division but could be summarized generally: to overcome a big problem existing for long in one field; to create a new research direction, new discovery or theory or interpretative framework; to breakthrough in key technology; to offer systemic solutions; to successful technology transfer for major societal or economic effects; to offer significant influential consulting suggestions (Bai, 2013). The newly categorized four groups target at long-term mainstreamed development of CAS by joint effort of institutes rather than the separated and scattered

short-term policies and projects. The future distribution of research funding will strengthen the potential major outcomes in more focused research areas. Besides, a mechanism for convergent investment of human resources, funding, and facilities is in construction to increase the efficiency. The principle was to encourage scientists to collaborate on fewer but large projects, rather than disparate and small projects mainly for seeking multiple grants.

5. Typology of Governance Conflicts around Strategies, Funding and Operation

As the fundamental causes of the reforms, tensions have existed for many years and hindered governance effectiveness. They may not be easily and completely solved by the reforms but at least they were realized and targeted to be mitigated or buffered. This section summarizes the substantial or potential types of conflicts of governance in CAS from key stakeholders, revealed by policy documents and interviews. Table 1, the three-by-three matrix, categorises the core conflicts at each level on each aspect. Figure 1 illustrates how the actors are generally involved for the following analysis level by level.

Table 1. Typology of conflicts aspects at three levels

Levels	Levels		
	C1: Strategy	C1: Strategy	C3: Operation
L1	Following political targets as public justification; Interpretation of the whole organizational strategy from L1 to L2	Allocation of institutional funds	One-size-fits-all regulations
L2	Autonomous development of legally independent institutes;	Stable institutional funding vs. competitive project funding	Cooperation and competition between institutes
L3		Exhausting fighting for competitive projects	Scientists' freedom

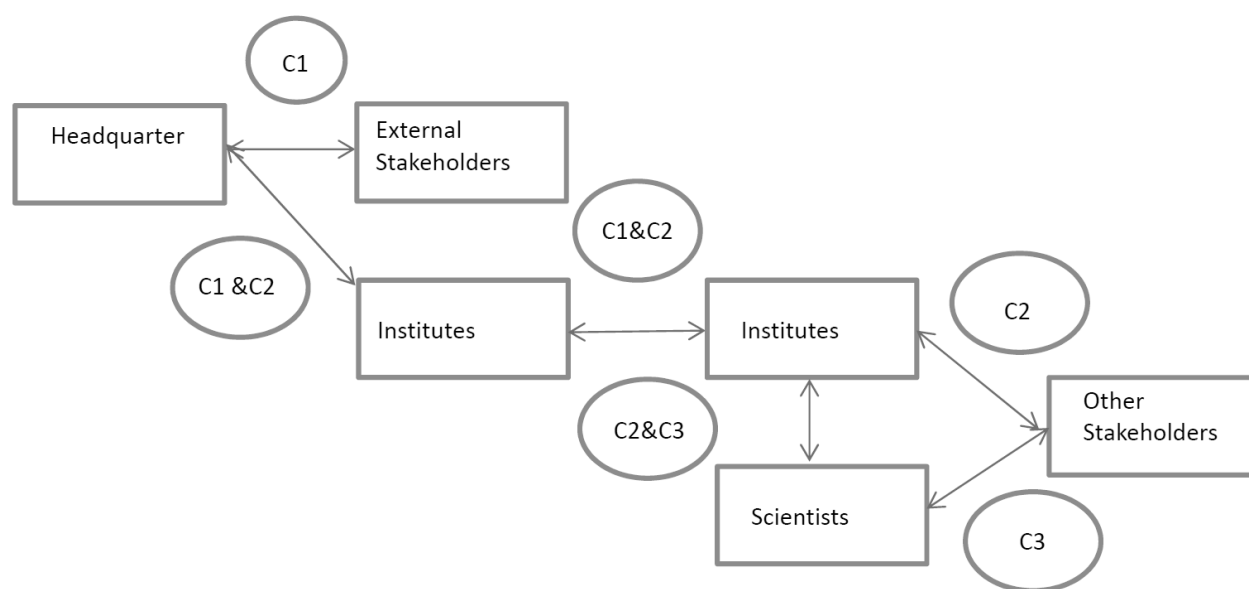


Figure 1. Stakeholders' involvement around the conflicts

5.1 Tensions between Headquarter and External Stakeholders

Within centrally planned system, a close entanglement of science and Communist party politics is easily set up (Good et al., 2015). Politics always give new elements so organizational missions have developed to mixed or complex ones over years (Crow & Bozeman, 1998). Within China's national innovation system, the State Council distributes S&T tasks through national policies, like Five-Year-Plan, to CAS and Ministries mostly in forms of huge funding programmes. MOST is the most important partner of CAS to jointly accomplish these programmes also as the major competitive resources for CAS institutes. The Ministry of Finance is in charge of block public finance of CAS. Ministry of Education guides the talents training of CAS universities. In general, CAS does not face many external stakeholders from wide-arranged background at L1 (like academics and business). The other external actors: regional governments as partly funders, Research Funding Council, scientific community, universities, industries, only collaborating

in joint programmes but do not impact organizational governance of CAS at all. "The whole governance is totally internal issue for CAS" (interview of Institute Director, 2013). However, China's leading universities and other research institutes are increasingly competing with CAS. There is debates about the whole position and whether CAS should stay such super large scale (Cyranoski, 2014). Therefore, the external justification of CAS has become urgent.

Additionally, the scientific responses of CAS as a whole to national strategic demands did not consider institutes' scientific responses in a bottom-up way. The diverse research orientations of CAS do not fit in one-size-fit-all requirements. The regular interactions between the State Council and CAS take place only by high-level leaders at national political conferences, in terms of annual performance reports of CAS as ex-post justification, and negotiation about new policy-making as ex-ante policy interventions to future plans. The intermediate monitor is mainly taken by financial audit. Here the political targets as overarching guidelines are pushed in typical

top-down ways from the State Council to CAS Headquarter, then to member institutes. Very little bottom-up scientific response and feedback under various research orientations are taken account into internal policies. To take one example, economic impact of R&D has been emphasized since 1990s when technological market was dominated by a few active enterprises. Especially after ‘Economic Crises’ in 2008, technology transfer capacities of CAS institutes have been invested through national ‘Pioneer Constructions’ with huge inputs of time, talents and infrastructures. CAS encourages all types of innovative technologies to be applied in products as many as possible, and push practical application as one-size request for all institutes. Meanwhile, most Chinese enterprises cooperating with CAS are not highly capable of R&D, usually expecting quick return on the short-term investment (two to three years) and without taking high risks. As for some inevitable longer-term risks like natural resource related R&D, if enterprises are not able or would not like to take, or if the market is very negative, the MOST and state-owned companies would take over with national support. But technological market always changes very fast. So are the relevant national and CAS policies. Such instability, for instance in every Five-Year-Plan, are inconsistency of research and strategy of CAS which should steadily follow scientific disciplines and patterns. It is also one significant reason of the over complex research orientations from massive projects driven by short-term policies. A large number of institutes conducting pure fundamental research were stressed with heavy burden of technological requirements, and forced to expand project-driven branches which distracted a lot of resources from the main research orientation and strengths.

Quotes from Institute A (with many research orientations): it is the requirement of CAS that your research needs to consider economic and societal impact, and at best into market for people’s life

quality. This is the political requirement from central government by National Conference. We need the companies to do the “D” part after we have done the “R” part, under the practical context of industries and businesses.

Quotes from Institute B (superior technological innovation): the ability of the whole CAS to handle S&T investment risks was still weak and even the attitudes towards the risks were negative and conservative. We only dare to invest in those short-term promising fields, and somehow ignore the long-term potentially important but risky fields.

Quotes from Institute C (fundamental research as core strength): technology R&D part is increasing due to the policy encouragement. But it depends on your own position and strengths. Actually there is impact towards relevant funding. Policy is changing all the time, but you have to stick to your own orientation and uniqueness and expand them.

5.2 Tensions between Headquarter and Institutes

Quite understandably, facing the complex missions, research orientations and multi-channel funding framework, the core challenge for Headquarter is to govern the various institutes with heterogeneous R&D. Besides massive institutes, a variety of working units at L2 make the governance in unified approaches nearly impossible. Generally, complying with overarching policies of Headquarter is the principle of CAS institutes. Such alignment has been forced by annual administrative reports in the past decades. These reports were treated as regular and reliable resources covering comprehensively all previous accomplishment and future plans. Categories of research orientations (fundamental knowledge extension, strategic innovation, technology transfer, and economic and societal contribution) could be recognized in statistic dimensions rather than policy targeted objects. Additionally, these reports depended heavily on quantitative indicators (like publication) which

actually drove significantly visible productivities of CAS. But since KIP and Innovation 2020, strategic targets of the whole CAS emerged aiming at leading role within national innovation system and international communities. Taking advantage of potential synergies among institutes based on qualitative assessment was urgently required. Workshops and conferences are encouraged actively for more peer communications. Quota restriction of projects application has been used to weaken homogeneous competition among institutes. Yet, institutes may prefer short-term strategies to solve immediate problems (Good et al., 2015), especially with high proportion of third-party funding. Due to the uncertainty about future development regarding several reforms, more thought-out coping strategies are difficult for institutes. It becomes a contradiction between Headquarter's standardized strategy-making and institutes' autonomous development.

Financially, the fragmentation of research orientation is intensified by scattered project funds, and vice versa. Multiple budgetary channels mean varying requirements from different funding bodies, leading to the instability, unpredictability, and discontinuity of R&D governance between actors at L1 and L2. Crosswise projects have obvious discrepancy in terms of scientific results and market effects between scientists and enterprises. It is quoted that "only one third of long-term stable fund is from CAS treated as small stipend. We have to rely on competitive research programmes even though we know they are very scattered and short-term positioned" (interview of researchers, 2013). Interviews demonstrate some institutes had been proud to develop comprehensively until they realized themselves the challenge of limitedly increasing resources being negatively fragmented. Now governance reforms devote for increased stable institutional funds but it is still in question how to prioritize resources allocation as it is impossible

to fulfill this vision for each institute at present. The patronage-, merit- or metrics-basis become heated debates. Reform often implies re-allocation of resources according to new rules. It inevitably brings discontent and pressure to the person in charge (Poo & Wang, 2015).

Self-operation of institutes is still monitored by Headquarter through controlling quotes of project, recruitment and block budget. It matters significantly whether one institute is legally independent or not. "If yes, you have much more right to develop yourself, getting more fund, more projects, attracting more talents, making your own financial and personnel regulations (with guidance from Academic Divisions), otherwise you only strictly follow Headquarter rules" (interview of Institute Director, 2013). Therefore, legally dependent institutes rely on the affiliated center financially and operationally, under quantity limit of program application, tenured staff, students educated etc.

5.3 Tensions between Institutes

At the meso-level, the research heterogeneity can be attributed to unordered competition and lack of incentives for communication and coordination between institutes. Balance between competition and cooperation turns to be coordinated between newly categorized clusters. They were isolated from each other due to geographic separation and cross-institutional barriers. The regional analysis result of KIP (Zhang et al., 2011) showed institutes in Beijing and Shanghai performing better than institutes in other cities during the same period. It is not uncommon that researchers across different institutes working on related projects but barely know each other. "There are nearly 20 institutes in biological sciences loaded with low-quality programs that are often repetitious and fragmented. There are more than 20 institutes with research program on LED isolated from each other, making industrial and commercial translation ineffective" (interview,

2013). Even within one complex structured research centre, occurred both duplicating work and missing opportunities to share knowledge, and unexpected competition between thematically different groups with external collaborators. Now teamwork is the keyword (Poo & Wang, 2015). The four categories of the PAP create environment where scientists selected under each group have to work together. Stable research funding and additional salary bonus should become incentives for team members of the new units to relieve their financial anxiety. Corresponding policies on authorships in publications and credit allocation of intellectual properties will be established to facilitate collaboration and teamwork. Not surprisingly, it would be a long exploratory way because of the long-standing characteristics, differentiated structures and newly developing research orientations.

5.4 Tensions between Institutes and Scientists

Bureaucratic burden of scientists weaken their creativity and productivity, especially the exhausting fighting procedures for various projects. Financially, shortage of block funds are regarded by researchers in a very unfavorable light which makes institutional funding unreliable and challenging for future planning. Anticipated responses towards newly classified research orientations are mostly from leading talents who are able to construct new groups for promising breakthroughs with supportive policies. 'It is important to have institutional money in order to fund development. Without stability, you quickly lose the best people' (Good etc., 2015). The CAS scientists' total income was far below the level of those in western countries, especially early-career researchers and PhD students. It is more difficult for them to produce indigenous innovation without abundant research funds and scientific networking. Besides, there is another monetary tension between personnel and non-labor expenses (facilities and infrastructures etc.) in research programmes. Many

researchers interviewed worried the usage efficiency of the costly infrastructures was very low. The labor expenses is strictly limited, usually below 30% of the total budget of programmes. "In fact, it is always under 15% of personnel costs in projects expenditures and these payments could only go to assistant students or temporary contract worker rather than official staff whose salary are considered covered already by Ministry of Finance" (interview of HR Departments of institutes, 2013). As for the increasing R&D funds of CAS, only about 25% of them were put into personnel but over 70% into importing expensive large-scale infrastructures and other facilities (laboratory equipment and experimental reagents) which "seemed more reasonable" (interview of researchers, 2013). There are particular procedures of their application, assessment, and appraisal. They are eventually funded by the Ministry of Finance, which is treated as making up for shortage of stable funding.

6. Evaluation Mechanism Dynamics

This section is about how evaluation mechanism works as governance instrument to cope with those tensions. We define evaluation mechanism as the ways the whole evaluation system is constructed and works, involving but not specifying all types of evaluation activities: institute, programme and project, talents, and labs respectively. Each of them has individual aims, rules, contents, and procedures closely in tune with organizational missions under particular contexts. For now there is still no external evaluation of the whole CAS yet, only big overarching programmes like KIP involving most of the institutes. The internal evaluation mechanisms have experienced decades of learning process to be integrated and mainstreamed in the organizational governance. We illustrate the two main functioning approaches in such learning processes. The first is institutionalization process of evaluation with four

identified distinct stages (Li, Shi, & Zhou, 2009) (Dai, Xiao, & Li, 2014). And the second is the stakeholders' communication and interaction patterns taking place within evaluation mechanism. They both reflect the addressed importance of R&D evaluation and the growing evaluation culture in CAS. It should be noted that the confidentiality of CAS evaluation is very high, neither guidelines nor any kind of result is next to impossible for external access. Awareness of transparency as justifying public funding is still very weak in China.

6.1 Institutionalization of Evaluation

Institutes' reports at the annual assembly of CAS have been constantly working as the most normalized approach for Headquarter's predominance. Funding-oriented projects, the major manner of conducting research in CAS, are evaluated to guarantees research quality with typical procedures of ex-ante appraisal in application, intermediate assessment, and ex-post check. However, institute evaluation has been increasingly promoted to a core and highly institutionalized concept as a governance tool by the reforms through four stages (Li et al., 2009) (Dai et al., 2014). Before 1993, there was no regulation on frequency or indicators of evaluation, no impact at all. During the first period from 1993 to 1998, it started with comparison of quantitative outputs from institutes like publications, patents, awards, title promotion, still without much consideration of impact. Then from 1998 to 2004 during the KIP, appeared institute ranking by scores from reviewers, mostly domestic but a few international, which provide explicitly evidence for governance decision-making. The third stage was comprehensive evaluation (2005-2010) with strong policy orientation and increased focus on institutes' creativity, but only based on institutes' initiative self-evaluation. Since 2011, "One-Three-Five" Institute Evaluation (One Position, Three Breakthroughs and Five Key Fostering Directions)

as a new mode began to drive more mature evaluation culture and strategic use, being institutionalized with a few pilots (four in 2012 and fifteen in 2013) but targeting at all institutes before 2020.

The shift to this dominated "One-Three-Five" evaluation occurred in the context of the 2012 Reform of Major Outcome Orientation to fulfill targets in "Innovation 2020". Position refers to institutes' current status and expected strengths and goals in terms of characteristics and ranking in upper or lower echelon within scientific communities nationally and internationally. Breakthroughs mean the most innovative or valuable contributions from both previous accomplishments and future short-term and longer-term targets. Key Fostering Directions involve the working themes and training groups accumulating the major resources with potential for maintaining and improving Position and Breakthroughs. All these "One-Three-Five" aspects compose the Major Outcomes of institutes which are ambitious to comprehensively summarize all the existing evaluand (projects, talents, trainings) with wide-range and long-term considerations. Each institute has signed its tasks contract with Headquarter during the Eleventh 'Five Year Plan'. And there would be strict implementation of expert diagnosis on "One-Three-Five" aspects every five years based on annual supervisory indicators (Evaluation Research Center, 2013). The concept of 'strategic evaluation' has been emphasized to improve evaluation at all levels by setting Major Outcome orientation rather than concrete standards for assessing research projects or policies. In this way it stresses the importance of respecting the differences between fields. This was not only a shift to a strategic evaluation system but also sought to simplify, modernize and improve the whole governance system of CAS. The core purpose of evaluation became to stimulate and guarantee major R&D outcomes from institutes under the control of Headquarter. This principle has penetrated into

individual and clusters of institutes, with higher-level strategic targets beyond scattered and purely quantitative outcomes.

It is a reflexivity process of institutes when selecting these aspects with focus on future achievements. And these requirements simplify previous administration from Headquarter which now only supervises the key elements of Major Outcomes rather than comprehensive performances. Furthermore, these aspects required to be diagnosed for institutes' further improvement could only be accomplished by high-level peer experts. Meanwhile, more and more international experts are invited to participate in CAS's evaluations with their global standards and experiences. "They are high qualified experts, open-minded, critical, with similar research orientation, and working in the very same field" (interview of Evaluation Research Center and institute researchers, 2014). Additionally, the experts guidelines with varying research orientations reflect the principle of categorized evaluation. For instance, those experts in frontier and fundamental research are mainly international peers with outstanding reputation, while those experts in strategic high-tech or applied research also include domestic specialists and users with rich local experiences. Reviewers help institutes find out problems and make beneficial suggestions about their structure, management and outcome transfer.

6.2 Communication and Interaction Patterns

Institutional learning takes place through stakeholders' communication and interaction at each step of evaluation procedures, far beyond final results. Attention to and involvement of key stakeholders is presumed to enhance the design and implementation of evaluations and the usefulness of evaluation results in decision-making (Bryson, Patton, & Bowman, 2011). A series of "interactions between evaluators, key policymakers, managers, and staff" (Wholey, 1972) are required to ensure

that the evaluation mechanism develop to the expectations of the key stakeholders. The culture of learning from evaluation is characterized by a collective engagement identifying and addressing commonly held issues and initiatives with participation, dialogue, deliberation, and self-determination (Preskill & Torres, 1999). However, stakeholders' communication and interaction, both inside and outside CAS, both crosswise at the same level and the longitudinal between different levels are much difficult to trace than that of western countries. This has been proved by the separated administration work among Headquarter Bureaus, and the lack of coordination in research among the institutes working in similar fields. The ex-post evaluation of KIP in 2010 was the pioneer to publicly release the results of governance reform of CAS as responses to political requirement of the State Council. Since then, evaluation reports became updated legitimization tool with new interaction forms. Afterwards scientific responses towards national policy-making and policy intervention become more and more through the reforms. In 2012 Major-outcome orientation and the "One-Three-Five" evaluation was initiated as responses to the national policy "Deepening the Reform of the S&T System and Accelerating the Development of National Innovation System". Different reform policies have evolved through different stages and supported by the accordingly policy evaluations. Furthermore, between CAS and other external stakeholders, to compose and assess national huge strategic programmes, the "Chief Expertise" Review Panel serves as the highest level Advisory Board for consultancy, with scientists and Academicians from MOST, CAS, universities etc. Such top level political, strategic and scientific communication only takes place in very significant events. Not only highly strategic communication between the State Council and CAS, more evaluation reports oriented by Major Outcome would bring

another wave of political and scientific dialogues involving more actors.

Designed evaluation procedures require stakeholders to engage in mutual communication especially around diverging concerns. At the preparatory stage, the procedures could be about discussing and making guidelines, outlines, schedules and selecting reviewers etc. The establishment of “Experts Database” in evaluation needs to joint work by Headquarter and the evaluated institutes which are much more familiar with the specialized experts in their own fields. High-level scientific information are produced, transformed, negotiated and reported in intermediate processes (Kuhlmann & Edler, 1998). Afterwards, more bottom-up interpretation and recommendations from evaluation results take place at L2 and L3, and the feedbacks would be transformed to L1 strategically and financially. Within the whole process, Evaluation Research Center as a hub coordinates with multiple stakeholders, summarizes and interprets all the valuable evaluative information for decision-makers at L1 for the next-round improvement. In these ways, stakeholders’ engagement and considerations in discussing controversies together are promoted by institutionalization of intended evaluation procedures, as balancing activities vis-a-vis stakeholders. Meanwhile, some unintended or spontaneous activities inevitably happening, such as unofficial or personal discussions, may result in unintended change of their behaviors.

All of these patterns depend on the availability and quality of evaluation information and services which are strongly valued and kept in limited transparency in CAS. Both the regularities of behavior embedded in the contexts of CAS and the international scientific communities affects the communication patterns, the information communicated, the interaction undertaken and the efforts put into these activities. Some policy documents of CAS proposed to strengthen more

overlapping and interdisciplinary communication among the new categorized research groups and administrative Bureaus. We expect more related information open and transparent.

7. Roles of Evaluation Mechanism in Aggregated Governance

More efficient governance with mitigated conflicts is the blueprint of the series of CAS reforms. It has been and would be a long exploratory and learning process that evaluation mechanism performs as powerful governance instrument through the above two ways. Institutionalizing those procedures as an evaluation routine is understood primarily as a collective learning process and thereby increasing the capacity for self-management (Robbecke & Simon, 2002). So far the below roles are anticipated by the authors that evaluation mechanism has been playing at the three levels for more aggregated governance of the whole CAS. Some aspirations at L1 are explicitly from the planned effect of reforms documents, but main expectations at L2 and L3 are derived from interviewees’ perceptions. We also admit some conflicts could not be solved by evaluation and evaluation creates conflicts themselves, which are not addressed in this paper.

7.1 Focused Top-down Governance at L1

To maintain governance equilibrium at L1, Headquarter needs to moderate conflicts with external stakeholders and institutes. Firstly, the strict top-down governance from the State Council to CAS becomes a bit more interactive due to evaluative communications. Newly classified research groups (Centers of Excellence, Innovation Academies, Big Science Facilities, and Specialty Institutes) could interpret and respond to political targets differently according to their research orientations. Accordingly, corresponding instead of overarching policy evaluation has a steering role in arranging more

efficiently many sources of control and influence from Headquarter to targeted institutes. Categorized policy evaluation helps to improve research productivity and safeguard accountability with more accurate aiming to research excellence and application, research–industry cooperation, supply of human resources, and international research collaboration. The integration of research goals and stimulating collaboration and teamwork within each category and across groups has been strengthened. More flexible cooperation emerges like closer linkage of research and training between integrated groups and universities.

“One-Three-Five” Evaluation is a further step for simplifying complex governance of huge number of institutes. It would result in considerable stability and predictability of the CAS R&D system and make its strategic planning of the whole bulky organization possible. Headquarter would focus on the institutes’ long-term positional goals rather than fragmented ones driven by short-term projects. Discontinuities of various short-term or mediocre researches would be filtered by deciding the “One-Three-Five” aspects. Reflexivity of institutes themselves is also enhanced in this process. Aggregated R&D outputs with more synergies among institutes would be facilitated by taking advantage of “One-Three-Five” Evaluation. In this way, more centralized scientific evidence and information are produced to depersonalize and depoliticize decision-making on resources allocation at L1. It could be traced from policy documents implying an automatic link between research evaluation and allocation of increased institutional funding. To be a safeguard against nepotism and corruption, governance would enforce and expand the use of merit-based evaluation of R&D, especially through implementation of international reviews.

7.2 Increased Autonomy of Institutes at L2

The reforms actively push the governance of CAS institutes forward to a new mode with simplified

administration and fewer tasks with concentrated investment. Evaluation of KIP proved increased performance of accumulated centers as one legal entity than previous unmerged institutes (KIP Evaluation Report, 2012). Each legally independent unit has a certain degree of scientific autonomy and freedom of constructing R&D groups and internal resources allocation. Major Outcomes of these units are guaranteed by the latest “One-Three-Five” evaluation (including reporting, assessing and controlling) from L2 to L1. Meanwhile, scientific empowerment has been promoted from L1 to L2 which allows for more space of other R&D funded externally beyond “One-Three-Five” aspects, like third-party funded crosswise projects. Additionally, with clearer and focused mission, the legally independent institutes save some resources used for overweighed administrations like many bureaucratic procedures of Headquarter.

Autonomy means more flexible self-operation of institutes in many aspects: international collaboration, translation and transfer of knowledge and technologies, talents attraction and retention. CAS reforms have issued a lot of open policies to attract actively academic and industrial collaborators both domestically and abroad, such as an “International Recruitment Plan” with “1000 Talents Project” since 2008 and “10000 Talents Project” since 2012 for high-quality scientists and innovation platforms of international standard. Recruitment of leading talents has been a key to an improvement of productivity in CAS (Liu et al., 2011). But the actions of collaborations still depend extremely on the institutes’ own scientific communities and scientists’ personal networking. More and more international evaluations create far more communication opportunities for talents attraction and retention.

7.3 Protection of Freedom and Trust of Scientists at L3

The restructure of CAS increases research volume

and quality, and to narrow the gap between China and other developed countries. The results of KIP show higher financial support and prestige position reduced the brain drain of scientists in China (Liu et al., 2011; KIP Evaluation Report, 2012). More importantly, focused top-down governance and increased autonomy of institutes protect scientists' creativity and productivity.

Rigorous evaluation does not mean zero freedom. The balance between them is trust. Researchers at L3 participate in all evaluation activities although they do not like it, shown in our interviews. They elaborate organizational strategies into scientific details on the one hand and also interpret scientific evaluation results and recommendations for strategic purposes on the other hand. This two-way process in CAS is and will be experiencing less vigorous monitors but more learning opportunities from more strategic and open-minded international evaluations. Scientists' freedom and trust are protected to promise the continuity of "One-Three-Five" aspects in very long-term. Evaluation contributes to excellent reputation of institutes and the whole institution which in the long term improve scientists' trustworthiness. Moreover, increased institutional funds especially more proportion on personnel allows space and flexibility of scientists' freedom, which can be only given based on good evaluation results. Besides, the confidentiality of evaluation reports is to protect research itself as well as scientific independence and freedom at L2 and L3. And evaluation of education and training offers guidance for professional personal development including expertise, creativity, motivation and independence of the employees.

8. Conclusion

Overall, this article makes the proposition: aggregated governance of CAS is supported by the evaluation mechanism which considers, reflects and mitigates the conflicts of governance. Within national and institutional contexts, the conflicts could be

generalized in three aspects: strategy, funding and operation at the three levels of organizational framework: central supervisory body at the macro-level (L1), institutional performance at the meso-level (L2), and individual research at the micro-level (L3). We argue the evaluation mechanism involving all evaluation activities especially the newly developing "One-Three-Five" Evaluation with Major Outcome Orientation works as governance instrument through the institutionalization processes, and increased stakeholders' communication and interaction patterns by formal procedures and practical activities. During these learning processes, the CAS governance has been steering towards aggregation with more integrated R&D outputs and efficient allocation of resources at L1, increased autonomy of institutes at L2, and improved freedom and trust of scientists at L3. These anticipated roles of evaluation in organizational governance could inspire the following implementation of evaluation activities and provide insights for accomplishing the targets of the still ongoing reforms of CAS. Also other developing large-scale PRIs under similar organizational framework may learn both lessons from those tensions and experiences of implementing instruments especially big reforms on governance and evaluation. Our further research will explore more evaluation mechanisms of international PRIs (Shapira and Kuhlmann, 2015).

Acknowledgement

I would like to thank the China Scholarship Council (CSC) for its financial support of this PhD project. And the anonymous interviewees and reviewers are appreciated greatly for their participation and comments.

Note

The PhD project has three case studies, the other two from Germany: Max Planck Society and Helmholtz Association. Comparative analysis and insights are in another working paper.

References

- Amanatidou, E., Cunningham, P., Goek, A., & Garefi, I. (2014). Using Evaluation Research as a Means for Policy Analysis in a 'New' Mission-Oriented Policy Context. *Minerva*, 52(4), 419-438.
- Bai, C. (2013). S&T Evaluation Reform Orientated by Major Outcome CAS (Vol. 4, pp. 4).
- Borras, S., & Edler, J. (2014). Introduction: on governance, systems and change, *Handbook of the governance of socio-technical systems Explaining Change*. Edward Elgar: EU-SPRI.
- Bryson, J. M., Patton, M. Q., & Bowman, R. A. (2011). Working with evaluation stakeholders: A rationale, step-wise approach and toolkit. *Evaluation and Program Planning*, 34(1), 1-12.
- CAS, E. R. C. o. (2013). An introduction to Expert Diagnosis Assessment for CAS Institutes. Chinese Academy of Sciences.
- Crow, M., & Bozeman, B. (1998). *Limited by Design*: Colombia University Press.
- Cyranski, D. (2014). Chinese science gets mass transformation. *Nature*, 513(7519), 468-469.
- Dahler-Larsen, P. (2011). *Evaluation Society*: Stanford University Press.
- Dai, T., Xiao, X., & Li, X. (2014). The context-based selection of evaluation models: the case of Chinese Academy of Sciences (CAS). Paper presented at the 11th European Evaluation Society Biennial Conference, Dublin, Ireland.
- Edler, J., Berger, M., Dinges, M., & Goek, A. (2012). The practice of evaluation in innovation policy in Europe. *Research Evaluation*, 21(3), 167-182.
- Evaluation Research Center. (2013). *One-Three-Five Institute Evaluation of CAS*.
- Good, B., Vermeulen, N., Tiefenthaler, B., & Arnold, E. (2015). Counting quality? The Czech performance-based research funding system. *Research Evaluation*, 24(2), 15.
- Hojlund, S. (2014). Evaluation use in the organizational context - changing focus to improve theory. *Evaluation*, 20(1), 26-43.
- Kuhlmann, S. (1999). Evaluation of research policy as a moderation process. In V. Paces, L. Pivec & A. H. Teich (Eds.), *Science Evaluation and Its Management* (Vol. 28, pp. 8-16).
- Kuhlmann, S., & Edler, J. (1998). *Clearing House*. Helsinki Conference.
- Li, X., Shi, B., & Zhou, J. (2009). A study of the model and methodology for institute evaluation in CAS. Paper presented at the Evaluation Conference Orlando, Florida, USA.
- Liu, F.-c., Simon, D. F., Sun, Y.-t., & Cao, C. (2011). China's innovation policies: Evolution, institutional structure, and trajectory. *Research Policy*, 40(7), 917-931.
- March, J. G. (1991). Exploration and exploitation in organizational learning *Organization Science*, 2(1), 71-87.
- Michelle, G. (2015). Raising standards. *Nature*, 520, 13.
- Nowotny, H. (2014). Paper presented at the EU-SPRI Conference: Science and Innovation Policies: Dynamics, Challenges, Responsibility and Practices, Manchester, UK.
- Poo, M.-m., & Wang, L. (2015). on CAS reform and Pioneer Action Plan. *National Science Review*, 2(1), 12.
- Preskill, H., & Torres, R. T. (1999). Building capacity for organizational learning through evaluative inquiry. *Evaluation*, 5(1), 42-60.
- Report, K. E. (2012).
- Robbecke, M., & Simon, D. (2002). *The Assessment of Leibniz Institutes: The Relationship between External and Internal Evaluation*.
- Sanders, B. A. (2003). Maybe there's no such thing as a "good cop" - Organizational challenges in selecting quality officers. *Policing-an International Journal of Police Strategies & Management*, 26(2), 313-328.
- Shapira, Ph., Kuhlmann, S. (eds.) (2015): *科技政策评估来自美国与欧洲的经验 (Learning from Science and Technology Policy Evaluation: Experiences from the United States and Europe)*, Beijing/China: Scientific and Technical Documentation Press, ISBN : 9787502397876
- Simon, D., & Knie, A. (2013). Can evaluation contribute to the organizational development of academic institutions? An international comparison. *Evaluation*, 19(4), 402-418.
- Sun, Y. T., & Cao, C. (2014). Demystifying central government R&D spending in China. *Science*, 345(6200), 1006-1008. doi: 10.1126/science.1253479
- website, C. (2015). <http://www.cas.cn/zj/yk/>.
- website, Nature Publishing Index (2013) <http://www.natureasia.com/en/publishing-index/global/>
- Wholey, J. S. (1972). What can we actually get from program evaluation? *Policy Sciences*, 3(3), 361-369.
- Zhang, D., Banker, R. D., Li, X., & Liu, W. (2011). Performance impact of research policy at the Chinese Academy of Sciences. *Research Policy*, 40(6), 875-885.