11.30-12.45

Policies for Impact Evaluation

Kim Huijpen (chair) - VSNU

Jinwon Kang – Korea-EU research Center

Martin Szomszor – Institute for Scientific Information

#IoS20

AESIS

# Policies for Impact Evaluation

Sukiennice room

AESIS
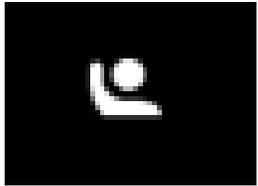
Broadcast permission:
- Turn on your microphone and/or camera
- Participate in the discussion
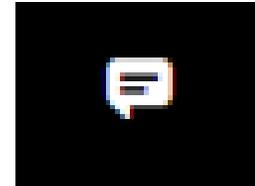
Conversations:
- General remarks
- Discussion
- News (links)
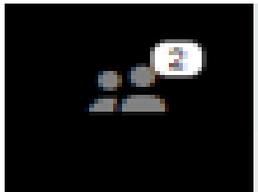
Who are the attendees?
- Speakers
- Participants

Q&A:
- (Targeted) questions
- Speakers answer the questions live

Lay out view:
Full screen, Tiled, Thumbnail

AESIS

#IoS20

Strategy Evaluation Protocol

2021–2027

VSNU KNAW NWO

AESIS Impact of Science virtual conference (Krakow)
4th of November 2020
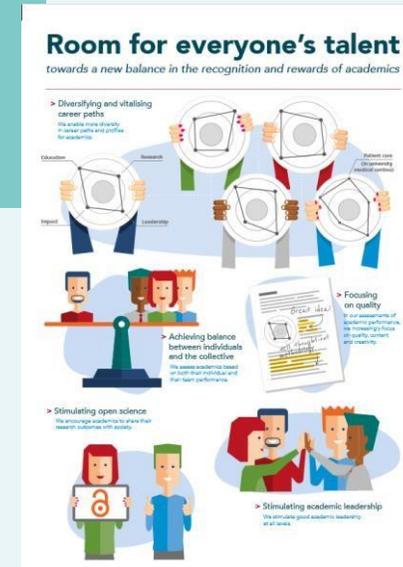
Kim Huijpen,
Programme Manager
Recognition & Rewards
The Netherlands
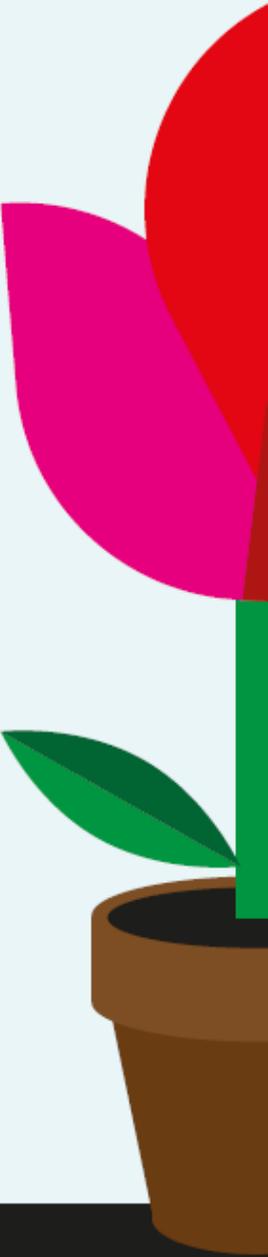
Background: SEP over the years

- SEP (since 2003/2009):
  - Substantive assessment that facilitates dialogue about research quality
  - The context/mission of the research unit is important
  - Societal relevance as a criterion (and a committee that can assess this)
  - Flexibility, also in indicators
- SEP 2016-2021: reflects, supports and drives developments
  - But: practice differs from the intention
- SEP 2021-2027: not just a text; **Standard** becomes **Strategy**
  - Additional explanations to promote that practice matches intentions

The Dutch context

Nov 2018	• Statement VSNU, NWO, NFU and ZonMw on Recognition and reward of academics
April 2019	• KNAW, NWO & ZonMw sign DORA (VSNU already did)
May 2019	• ZonMw & NWO conference Scientist 2030: Evolution or Revolution
Nov 2019	• Postion paper: Room for everyone's Talent • VSNU - EUA Conference on Recognition & Rewards
March 2020	• New Strategy Evaluation Protocol 2021-2027

Room for everyone's talent
towards a new balance in recognising and rewarding academics

Strategy Evaluation Protocol
2021–2027
VSNU KNAW NWO

Developing SEP 2021-2027

Meeting board members and deans & working conference

Consultation SEP 2021-2027

2018

Spring 2019

Summer 2019

Autumn 2019

Start 2020

Desk research & interviews by external party

Writing consultation version SEP

SEP approved by boards of VSNU, NWO, KNAW
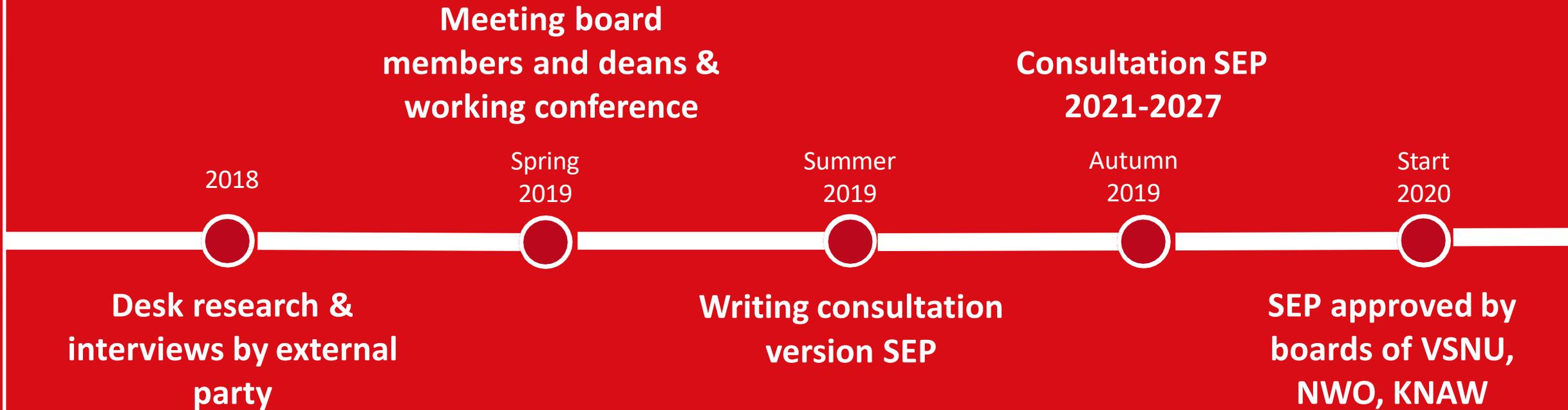
SEP committee

- Prof. T.T.M. (Thom) Palstra (chair)
- Prof. S.J. (Sijbrand) de Jong
- Prof. K.I. (Karen) van Oudenhoven-van der Zee
- Prof. F. (Frank) Miedema
- Prof. F.P.I.M. (Frank) van Vree
- K. (Kim) Huijpen, MSc (secretary)

SEP working group

- <u>Kim Huijpen (chair), VSNU</u>
- <u>Dov Ballak, NFU</u>
- <u>Chantal Bax, KNAW</u>
- <u>Dagmar Eleveld-Trancikova, Radboudumc</u>
- <u>Peter Hildering, QANU</u>
- <u>Lise Koote, VSNU</u>
- <u>Jacqueline Mout, NWO</u>

- Anne-Roos Renkema, VSNU
- Lambert Speelman, VSNU
- Els Swennen, Maastricht UMC+
- Haico te Kulve, University of Twente
- Leonie van Drooge, Rathenau Instituut
- Lieke van Fastenhout-Strating, University of Amsterdam

What remained the same? (1)

- Main goal is to maintain and improve the quality and societal relevance of research
- Assessment of a research unit is in light of its own aims and strategy
- A self-evaluation written by the unit forms the basis of the evaluation
- Evaluation is performed by an assessment committee appointed by the relevant board

What remained the same? (2)

- The SEP assessments help to monitor and improve the quality of research as part of the ongoing quality assurance cycle

- Additionally, the assessments of the research quality and societal relevance of research contribute to fulfil the duty of accountability towards government and society

- The assessment committee assesses the performance of the unit according to three main assessment criteria: 1) **research quality**, 2) **societal relevance** and 3) **viability**.

Most important changes (1)

- Bibliometric indicators are less important; self-evaluation is a narrative argument supported with factual evidence

- SEP explicitly follows guidelines of DORA

- No more 'quantitative scores' & rankings between research units but well-argued assessments with sharp, discerning texts

- Assessment committee weighs results and reflections of research unit on four specific aspects (three are new)

Most important changes (2)

- Assessment committee weighs results and reflections of research unit on four specific aspects (three are new):

- Open Science

- PhD Policy and Training

- Academic Culture and

- Human Resources Policy

Three criteria vs four specific aspects
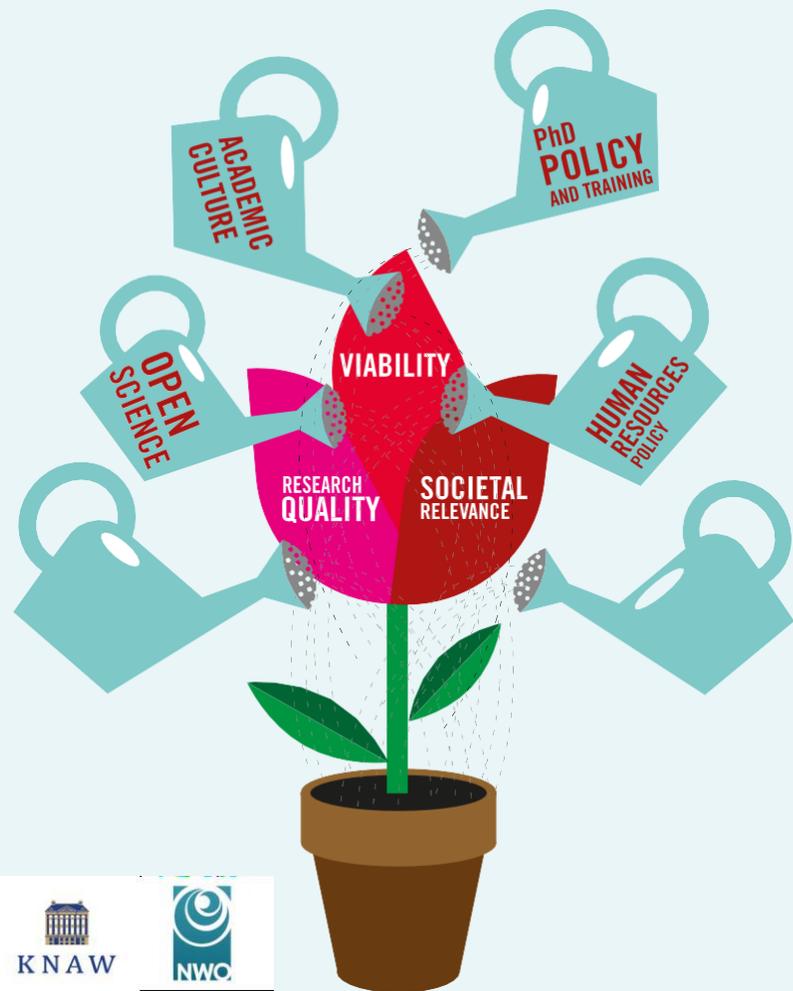
- Many possible aspects can be addressed when discussing the three criteria: these four are specifically highlighted in the SEP to ensure that they are reflected in every evaluation

- In principle not assessed separately (and described in the report), but as part of the discussion of the three criteria
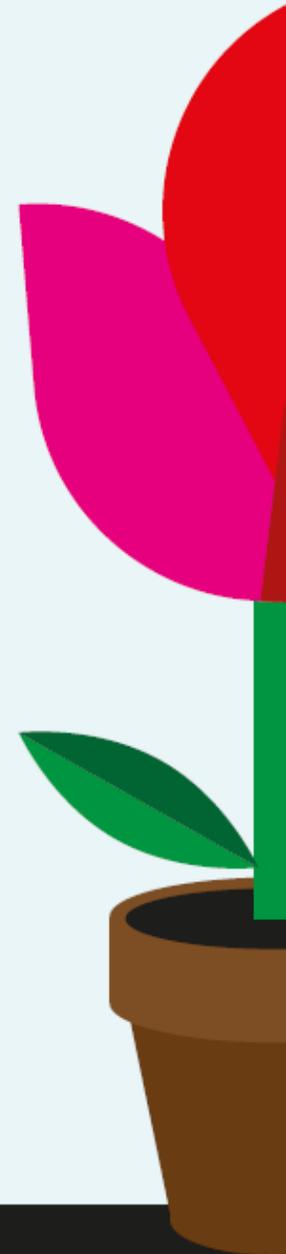
Specific aspects: Academic Culture

Openness, (social) safety and inclusivity:

- Appreciating multiplicity of perspectives and identities
- How leaders take responsibility to contribute to open culture

Research integrity:

- Unit's policy on research integrity and requirements Netherlands Code of Conduct for Research Integrity
- Reflecting on relevant dilemmas (e.g. authorship, privacy or collaborations with stakeholders)

Specific aspects: Open Science

The extent to which the research unit:

- Involves stakeholders and opens up its work to other researchers and societal stakeholders

- Stores research data according to FAIR principles and how it makes its research data, methods and materials available

- Makes its publications available through open access

- Pays attention to other aspects of Open Science

Specific aspects: PhD Policy and Training

- PhD programme content and structure

- Selection and admission procedures for PhD candidates,

- Duration, success rate, exit numbers and career prospects for PhD candidates

- Supervision, mentoring and coaching of PhD candidates

- Position of PhD candidates and PhD training in the unit's research

- PhD education at relevant institutional graduate school(s) and (national) research school(s) and its quality assurance system

Specific aspects: Talent Management

Human Resources Policy:

• Diversity in relation to aims, strategy and policy of research unit

• HR practices such as inclusive selection and appraisal procedures

Talent Management:

• Talent selection and development in relation to aims and strategy

• Unit's recruitment policies, opportunities for training and development, coaching and mentoring

• Properly evaluating, rewarding and incentivizing staff

Position paper

In November 2019, the Dutch Universities published the position paper 'Room for everyone's talent' together with Dutch public knowledge institutions and funders of research (VSNU, NFU, KNAW, NWO and ZonMw)

Why do we need a change in recognition and rewards?

Why a change is needed

What we want to see as a result

- A healthy and inspiring working environment.

- Scientists are valued for their particular talent, be it education, research, leadership, impact or patient care. Careers are possible on the basis of each of these activities, rather than research only.

- People are enjoying their work and are no longer stressed by the pressure of producing publications, as research exposure through other channels will be possible and valued.

- Science is practiced in teams and all team members receive credit for the team result.

- And last but not least, the recognition that academics receive needs to reflect what society expects from science.

What do we want to change?

1. Enable **diversification and vitalisation of career paths**, thereby promoting excellence in each of the key areas (education, research, impact, leadership and patient care)
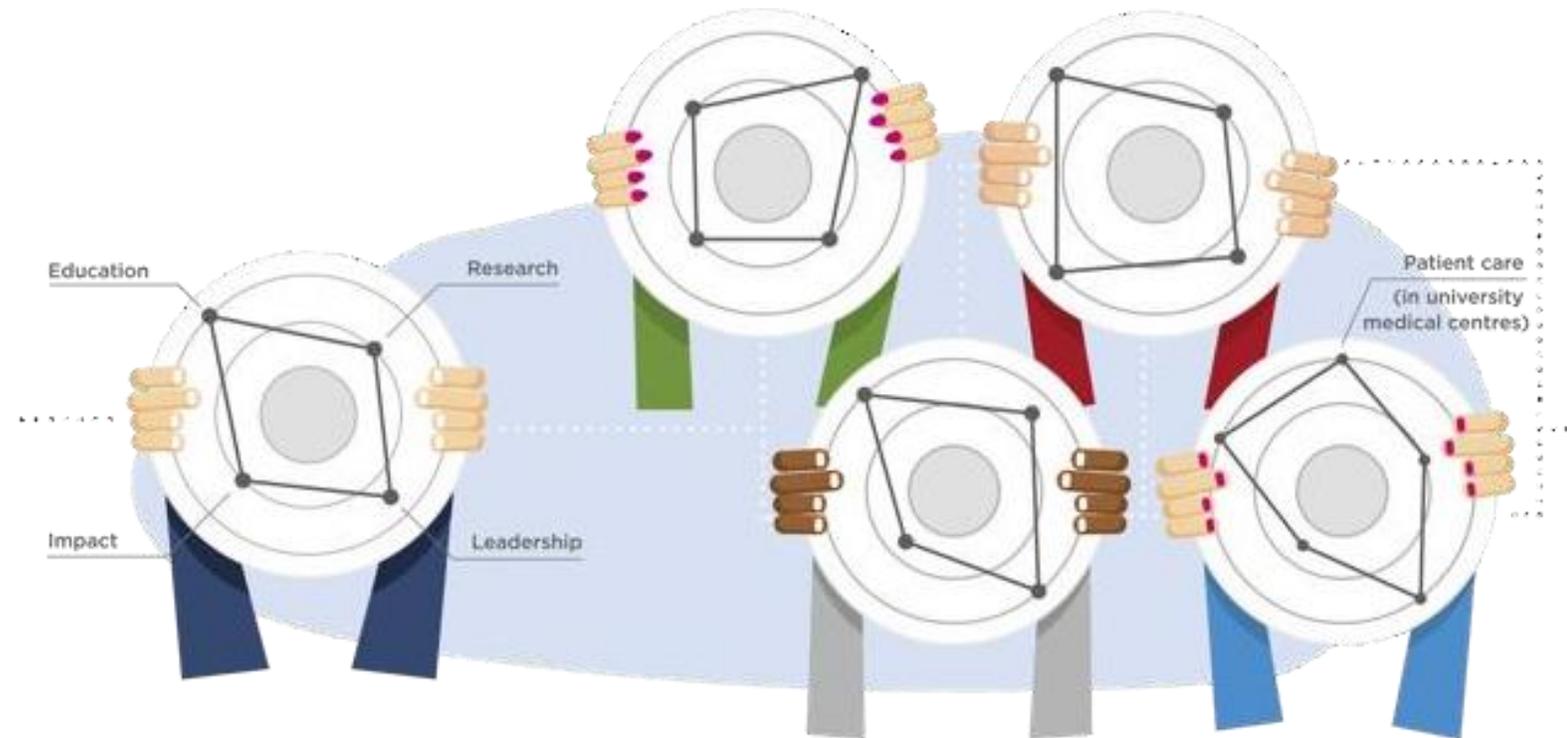
2. A better balance between individual and team performance

Inspire cooperation between organizations, disciplines and within teams (**Team Science**)

3. More focus on quality of work over quantitative results

Good scientific research increases scientific knowledge and makes a contribution to solving societal challenges

4. Open Science becomes the norm and stimulates interaction between scientists and society

Stimulating Open Science means recognizing and rewarding other aspects of research (in addition to publications), such as **datasets or software**, as important research outputs

5. More emphasis on the value of high quality leadership in academia to set the course in research and education, to achieve impact, and to ensure that teams of academics can do their work as well as possible

Let's move together!

Thank you for your attention!

More information: Kim Huijpen, Programme Manager
huijpen@vsnu.nl

Some interesting references

- Strategy Evaluation Protocol (SEP) 2021 – 2027
- Video Strategy Evaluation Protocol (SEP) 2021-2027
- Position paper 'Room for everyone's talent: towards a new balance in the recognition and rewards for academics',
- Youtube playlist Recognition & rewards
- Youtube video on our Recognition & rewards programme

- Illustrations by GREATGRAPHIC and Things to Make and Do

R&D Evaluation in Korea

2020.11.4

Jinwon Kang

Korea-EU Research Centre/KISTEP

Contents

1. **Background**

2. **R&D Evaluation System**

    2-1. Self/Meta Evaluation

    2.2. In-depth Evaluation

3. **Policy Issues**

1. Background

2. R&D Evaluation System

2-1. Self/Meta Evaluation

2.2. In-depth Evaluation

3. Policy Issues

S&T Environment in Korea

|	1960s	1970s	1980s	1990s	2000s
Policy Trends	**Industry**-oriented policy			**Technology**-oriented policy	
S&T Policy Direction	**Building R&D Infrastructure** ▮ GRIs established (e.g. KIST(1966)) ▮ MOST established in 1967 Daedeok Science Town built in 1974	**Promoting R&D** ▮ Enforcing the national R&D programs ▮ Promoting the university -based researches (e.g. SRC, ERC)		**Enhancing Technology Innovation** ▮ KISTEP(since 1999) ▮ Increasing efficiency of R&D investment (coordination of ▮ S&T-related policies) Planning Total Roadmap	
Changes in Major Industry	Primary goods	Light industry goods	Light & heavy industry goods	Heavy industry goods & electronic products	Advanced Electronic & transport products

R&D Investment Trends

- **Ratio of total R&D investment is 4.55% of GDP$_{('17)}$: 1st among OECD countries**

- **4.55% is composed of private sector (3.43%) and government (1.12%)**

million US dollar

Total R&D investment
Government R&D investment
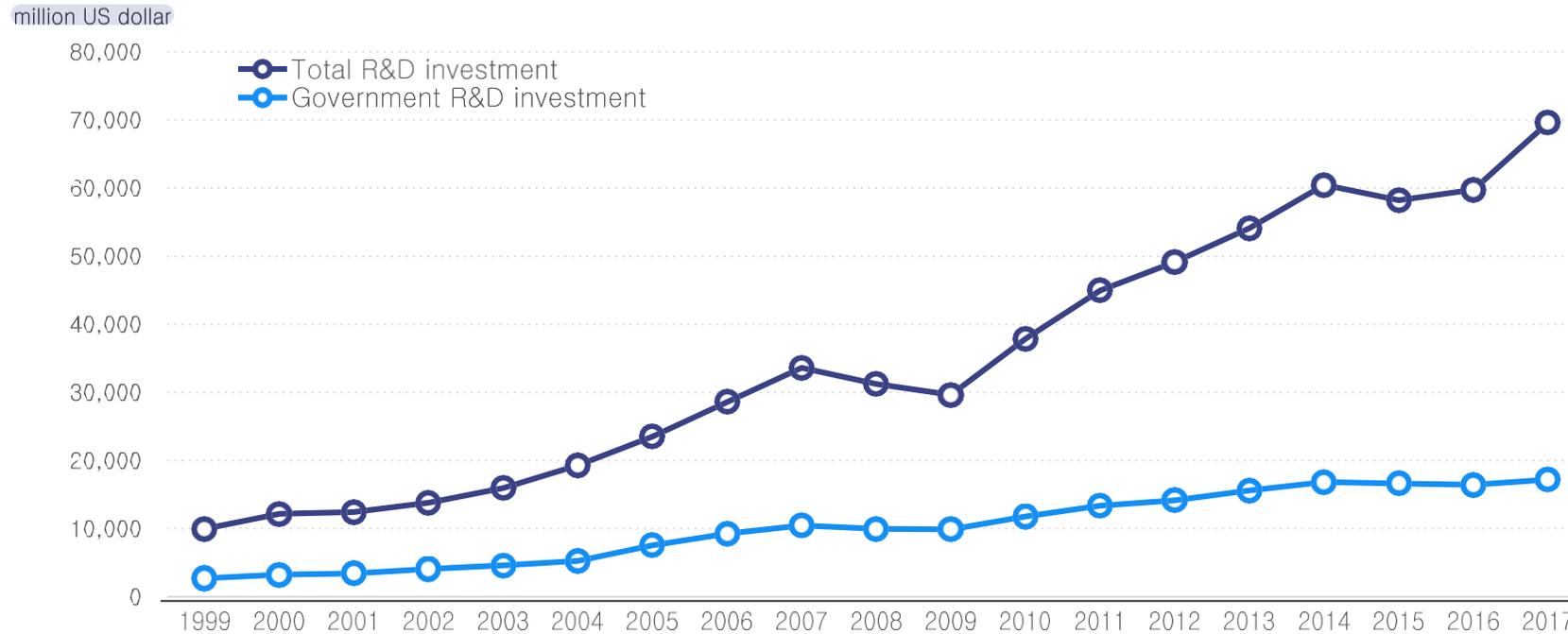
source MSIT, KISTEP (2018)

R&D Investment Trends

- **When government R&D budget amounts to 20 trillion Won(＇19), performance against investment became rising issue**

100 million Won

■ Government R&D budget (100 million Won)　—○— Variation over last year (%)　%

Year	Budget	Variation
2000	41,974	13.2%
2001	57,339	36.6%
2002	61,417	7.1%
2003	65,154	6.1%
2004	70,827	8.7%
2005	77,996	10.1%
2006	89,096	14.2%
2007	97,629	9.6%
2008	108,423	11.1%
2009	123,437	13.8%
2010	137,014	11.0%
2011	148,902	8.7%
2012	160,244	7.6%
2013	168,777	5.3%
2014	177,428	5.1%
2015	189,231	6.7%
2016	190,942	0.9%
2017	194,615	1.9%
2018	196,681	1.1%
2019	205,328	4.4%

source　KISTEP (2019)

Quantitative & Qualitative performance

Paper

No. of paper (SCI)

	55,800	60,475	63,311
	12th	12th	12th
	'14	'16	'18
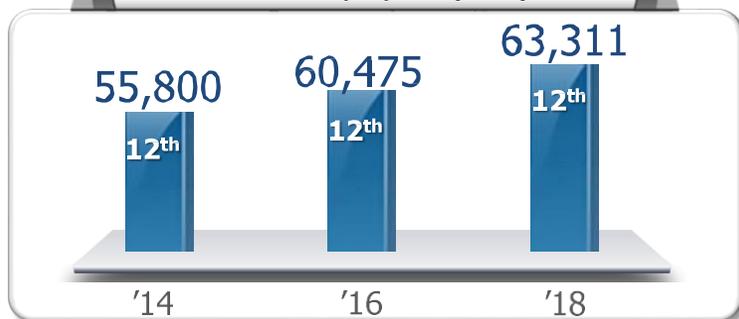
Patent

No. of patent (PCT)

1,460	1,955	2,330
'14	'16	'18
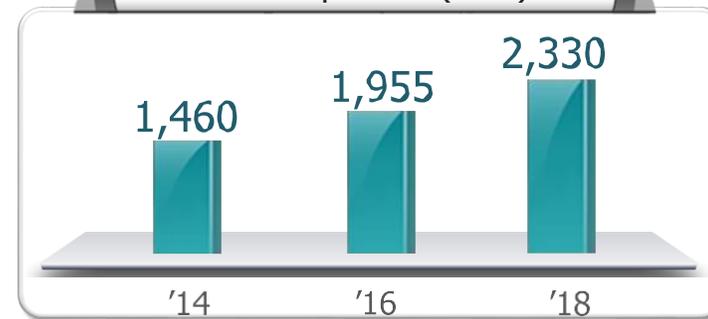
Paper

cited number per paper(ave.)

	4.79	5.09	6.36	6.04
	33th		33th	
	2009-2013		2014-2018	

- Korea(ave.)
- World(ave.)

Knowledge

Global Innovation Index(2020) (GII,WIPO)

7th	10th	15th
Knowledge Creation	GII	Knowledge Diffusion
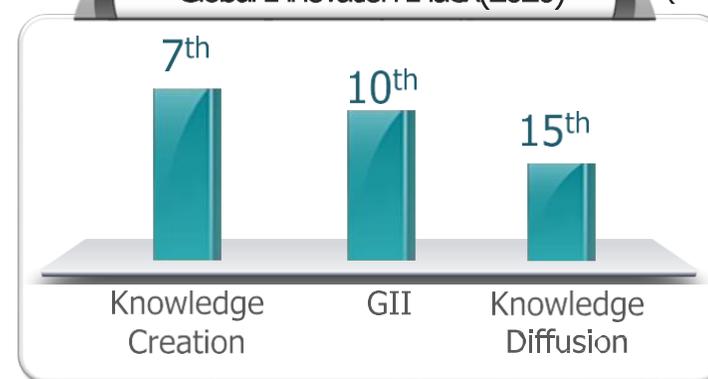
Toward leading from catching up in R&D

Catch up

▷ Easy to succeed

▷ Quantitative performance -centered

▷ Economic growth oriented R&D

▷ R&D by each technology and institution

Leading

▷ Creative, Challenging R&D

▷ Qualitative performance-based R&D

▷ Public and social problem solution R&D

▷ Fusion, cooperation R&D

1. Background

2. R&D Evaluation System

   2-1. Self/Meta Evaluation

   2.2. In-depth Evaluation

3. Policy Direction and Issues

Act to Evaluate Gov't R&D Performance

🔵 Enacted in Dec. 2005

- Establishing 5 year-based basic plan for R&D performance evaluation by MSIT (Article 5)

- **Setting goal, annual objectives and performance indicators for each R&D program** by gov't ministries (Article 6)
  - MSIP should develop and provide performance indicators available

- **Implementing in-depth evaluation and meta-evaluation** by MSIT (Article 7)

- **Implementing self-evaluation by line ministries** (Article 8)
  - Ministries are required to evaluate R&D programs except ones that are directly evaluated by MSIT (in-depth evaluation)
  - They should report the result of self-evaluation to MSIT (for Meta-evaluation)
  - * No obligation to submit the result of R&D project evaluation implemented by them

Type of Program Evaluation in Korea

- **Interim Evaluation**
  - Evaluates Each ongoing R&D program in terms of its performance every three years

- ❖ **In-depth Evaluation**
  - Is special interim evaluation, which is performed by KISTEP

- **Final Evaluation**
  - Evaluates the performance of a R&D program when it ends

- **Follow-up Evaluation**
  - Monitors technology transfer and commercialization after the program ends

Evaluation Unit

🔵 **S&T Policy**

- ◌ Spending Review is performed by MOSF (Not Evaluation)

🔵 **R&D program (847, '19)**

- Self/Meta / In-depth evaluation unit
- Self/Meta evaluation is organized by ministries and MSIT (KISTEP)

🔵 **R&D project (70,327, '19)**

- ◌ R&D project evaluation is organized by program-running ministries (35)

A Ministry · · ·

A Ministry

Program 1 · · · Program n

Project 1 · · · Project n   Project 1 · · · Project n

<R&D program/project structure>
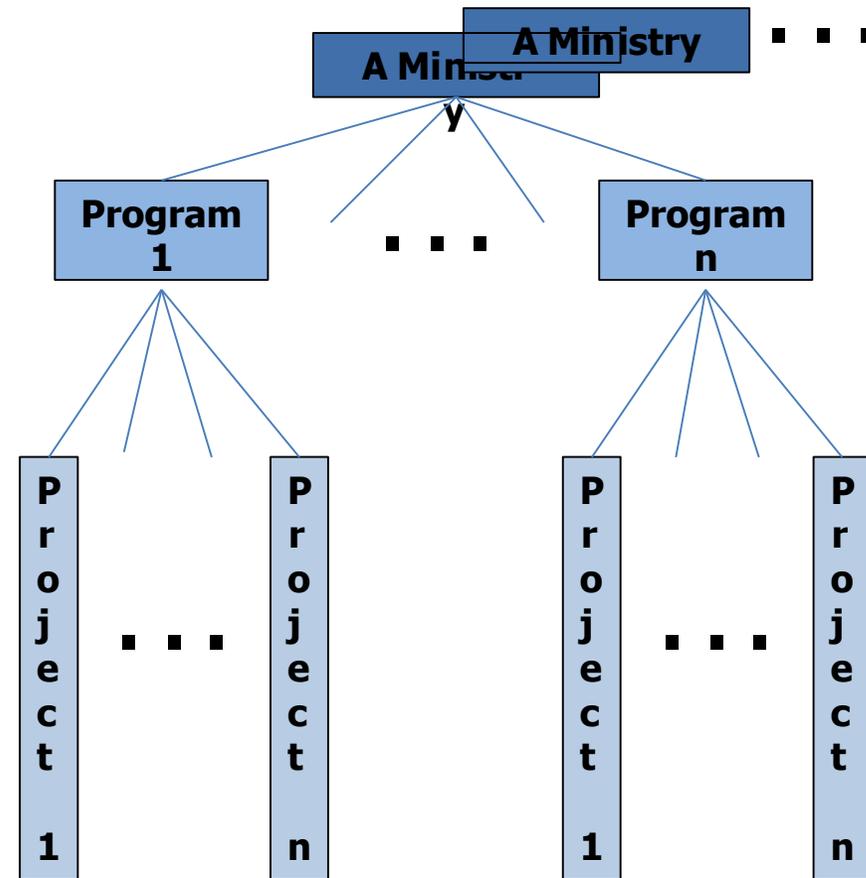
R&D Evaluation Structure

Characteristics of Self/Meta Evaluation

K-PART : benchmarked PART of USA

- Self Evaluation (Ministry) → Meta Evaluation (MSIT/KISTEP)

- Plan-Do-See Monitoring

- 3 year evaluation cycle (1/3 programs per year)

- Yes or No questions of indicators with weight

- A grade system of five categories
  (excellent-good-moderate-unsatisfactory-poor)

- Budget increase for excellent and good programs and budget
 cut for unsatisfactory and poor programs

Conceptual Procedures


Self/Meta Evaluation

[Step 1. Planning]
Setting of
Performance Objective &
Indicator

[Step 2. Management]
Management of
Performance Objective &
Indicator

[Step 3. Evaluation]
Evaluation of
Objective Achievement

[Step 4. Feedback]
Program Improvement &
Budgeting Allocation

Conceptual
Procedures

Procedures of Self/Meta Evaluation

PI Setting	Self-Evaluation	Meta-Evaluation
● MSIT + Ministry	● Ministry (Department)	● MSIT/KISTEP

Setting of Performance indicators → Self-Eval. Committee → Self-Eval. Result → Submit to MSIT → Meta-Eval. Committee → Meta-Eval. Result → National S&T Council → Budget Planning, Allocation
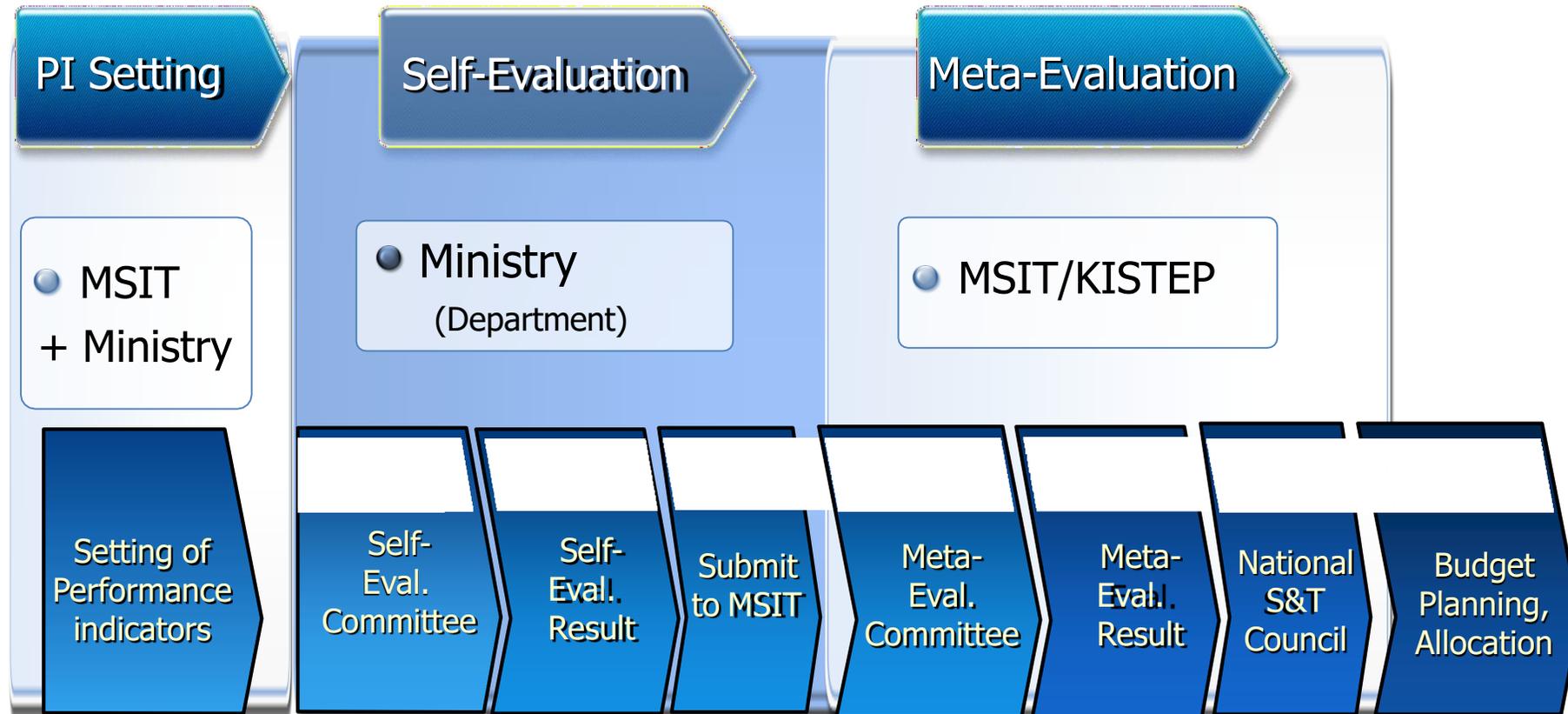
Performance Indicators (example)

Performance Indicators

| Scientific PI | Technological PI | Economic PI | **Social PI** | Infra PI |

Scientific PI
- 《Paper》
  ①SCI
  ②KCI
- 《Resource》
  ③Bio-Res.
  ④Compound

Technological PI
- 《IPR》
  ①Patent
  ②Other IPR
- 《Non IPR》
  ③Tech.Innovation
- 《Growth Engine》
  ④Contents/SW
  ⑤Service
  ⑥Product
  ⑦Plant
  ⑧Medicine/
  Medical Device

Economic PI
- 《Direct》
  ①Royalty
  ②Export/Import Substitution
- 《In-direct》
  ③Tech. utilization
  ④SME support
- ⑤ Tech.Comercialization
- ⑥R&D Service
- ⑦Employment (Private)

Social PI
- 《HR》
  ①HR Training
  ②Employment
- 《Region》
  ③Regional Development
- 《Welfare》
  ④Policy Effect (incl. Tech. Reg., Energy saving)
  ⑤Public Service
- 《Publicengagement》
  ⑥Promotion
  ⑦Diffusion
- ⑧International Cooperation (incl. Human Exchange)

Infra PI
- 《Infra》
  ①Sharing Infra
  ②Excl. Infra
  ③Comuterized System
- 《Space》
  ④Development
  ⑤Utilization
- 《Defence》
  ⑥Defence System

PI regarding type of program (example)

Type of Program	Short-term PI (Output)	Mid-term PI (Short-term Outcome)	Long-term PI (Long-term Outcome)
1. Basic Research	Sci/Tec	Sci/Tec	Sci/Tec/Eco/Soc
2. Short-term Industrial TD	Tec/Eco	Tec/Eco	Tec/Eco
3. Mid/Long-term Industrial TD	Sci/Tec	Sci/Tec/Eco/	Eco
4. Public TD	Sci/Tec	Sci/Tec/Soc	Eco/Soc
5. Regional R&D	Sci/Tec/Eco	Tec/Eco/Soc	Eco/Soc
6. Defence TD	Tec/Eco/Inf	Tec/Eco/Inf	Tec/Eco/Inf
7. HR Training	Sci/Tec	Sci/Tec	Eco/Soc
8. Equipment/Facility Building	Inf	Inf	Eco/Soc/Inf
9. Performance Diffusion	Tec/Eco/Inf	Tec/Eco/Inf	Tec/Eco/Inf
10. International Cooperation	Sci/Tec	Sci/Tec	Eco/Soc

Type of self-evaluation

- **Self-evaluation can be divided into Interim, Final, Follow-up evaluation** regarding the time to evaluate

  - **Follow-up evaluation** is performed 3~5 years after the program's completion

Interim Evaluation		Final Evaluation		Follow-up Evaluation
· Achievement of performance target (30%) · Excellence of performance (70%)	→	· Achievement of performance target (70%) · Plan of performance utilization (30%)	→	· Performance utilization system and activities (50%) · Utilization result and **ripple effect** (50%)
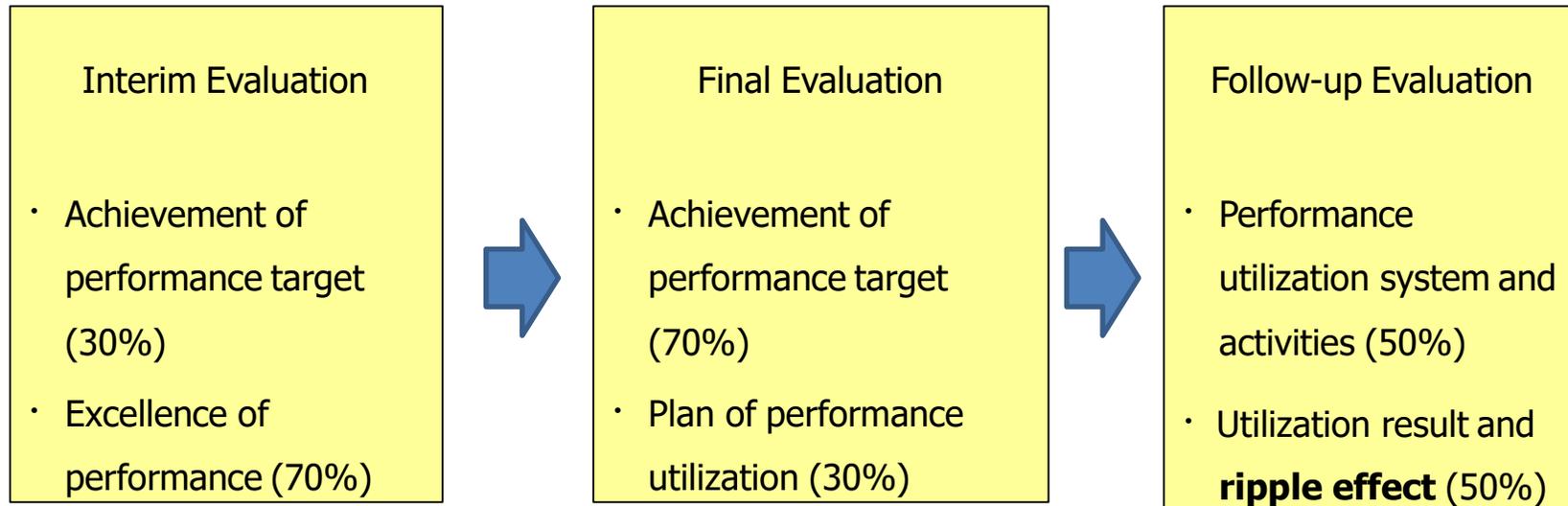
Evaluation scores using PI (Interim/Final)

Is performance achieved quantitatively as planned?

Indicator scores according to performance achievement of program.

Performance indicators	Weight (a)	Target (b)	Accomp. (c)	Accomp. rate (d=c/b)	Indicator appropriacy (e)	Indicator score (f=30xaxdxe)	Final score (g=∑f)
performance Indicator 1	0.4	100	110	100%	1	12	23.0 / 30
performance Indicator 2	0.4	150	103	69%	1	8.3	
performance Indicator 3	0.2	20	18	90%	0.5	2.7	

Analysis of Ripple effect (Follow-up)

Is the result utilization and its ripple effect excellent?

- (Scientific) Contribution to Basic Science

- (Technological) Improvement of Industrial Competitiveness

- (Economic) Creating Market/Employment or Increase of Sales
- (**Social**) Improvement of Security, Decrease of Energy Cost, Decease of Regional Gap

- (Infra) Sharing of Equipment/Facility

- ❖ The level of ripple effect: World best or similar level, Korea best or similar level, Korea average below level

Type of Assessment

- **<u>Evaluation:</u>** The systematic and objective assessment of an on-going or completed project or program, its design, implementation and result. The aim is to determine the relevance and fulfillment of objectives, development efficiency, effectiveness, **impact** and sustainability.

- **<u>Review</u>**: An assessment of the performance of an intervention, periodically or on an ad hoc basis. Reviews are usually less comprehensive and/or in-depth than evaluations. They tend to emphasize operational aspects

- **<u>Monitoring:</u>** A continuing function that uses systematic collection of data on specified **indicators** to provide management and the main stakeholders of an ongoing development intervention with indications of the extent of progress and achievement of objectives and progress in the use of allocated funds.

* OECD/DAC (2002), Glossary of Key Terms in Evaluation and Results Based Management

Criteria of program selection

- **Selection criteria of program for in-depth evaluation**

- Long term R&D programs with large budget

- R&D programs that are suspected to duplicate with each other

- R&D programs that are jointly implemented by a couple of gov't ministries

- R&D programs that are at issue

- R&D programs that MSIT judges are necessary to evaluate

Perspectives of in-depth Evaluation

- In-depth evaluation has been performed in accordance with 4 evaluation perspectives. In-depth evaluation is also focused on the Efficiency

Evaluation perspective	Evaluation contents
Relevance	• Compliance with higher level plan or strategy / relevance of strategic objective
	• Relevance of investment strategy
	• Appropriacy of government R&D support
Effectiveness	• Achievement over objective
	• Economic and social effect
Efficiency	• Efficiency of performance
	- Output over input analysis
	- Qualitative analysis of excellent performance
Systematic nature	• Performance management and utilization
	• Possibility of overlapping and the necessity of collaboration

The Evaluation Framework

Needs	Inputs- Implementation - Outputs	Short-term Outcome	Long-term Outcome

Co-authorship of researcher Reciprocal & strategic collaboration	Sharing Research Resources (Money, information and HR)	Co-performance (paper, patent…)	Collaborative Network for joint research	(Improvement of Research Capacity)

1. Is the investment by the programs are relevant to the policy priorities?

2. Are the program implementing structures are efficient ?

3. Do the programs achieve the objective?

<Relevance >   **<Efficiency >**   **<Effectiveness >**

Relevance

- **Is the cooperation is relevant to the policy direction ?**
  - Reciprocality of international joint research in terms of objective and sharing resources (funding, information, infrastructure, HR)

Reciprocality of objective

0%
0%
8%
10%
25%
57%

- ① Introduction of advanced technology
- ② Production of global level performance
- ③ Utilization of foreign research facility
- ④ Matching fund with foreign researchers
- ⑤ Overseas market expansion
- ⑥ Etc

Reciprocality of sharing resources

1%
16%
32%
19%
16%
16%

- ① Research funding
- ② Research information
- ③ Research facility
- ④ Researcher
- ⑤ Research material
- ⑥ Etc
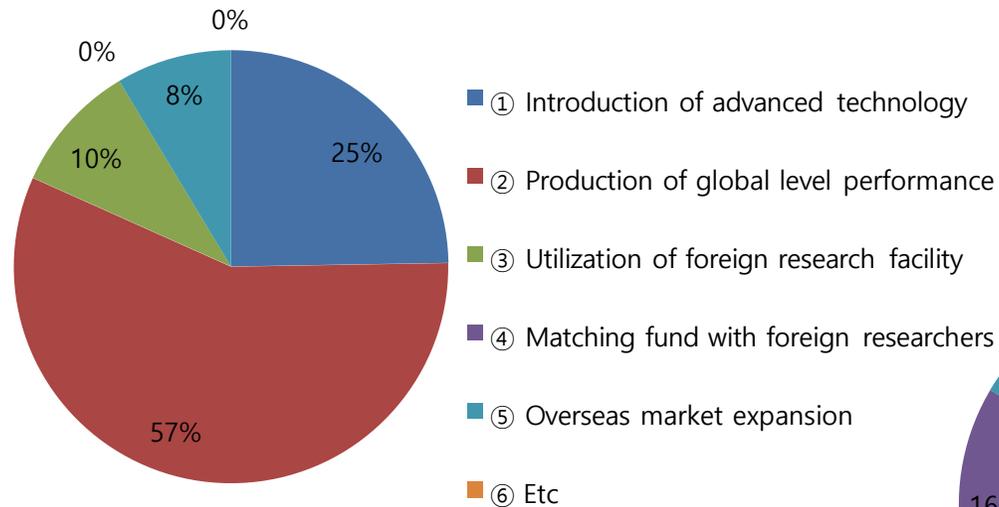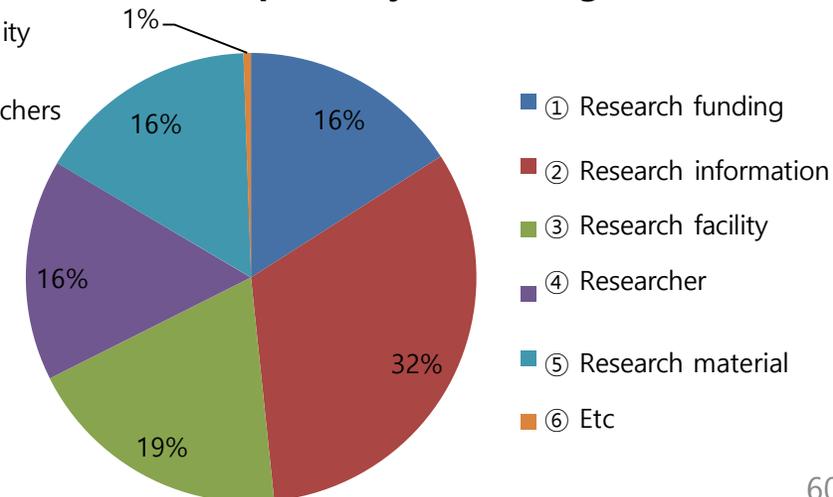
A Case: International joint research programs

Relevance

- **Is the cooperation is relevant to the policy direction ?**

- Strategic cooperation of international joint research in terms of research field and

  cooperation country

Strategic cooperation

- ① Bottom-up of field and country
- ② Top-down of field, bottom-up of country
- ③ Bottom-up of field, top-down of country
- ④ Top-down of field and country
- ⑤ Etc
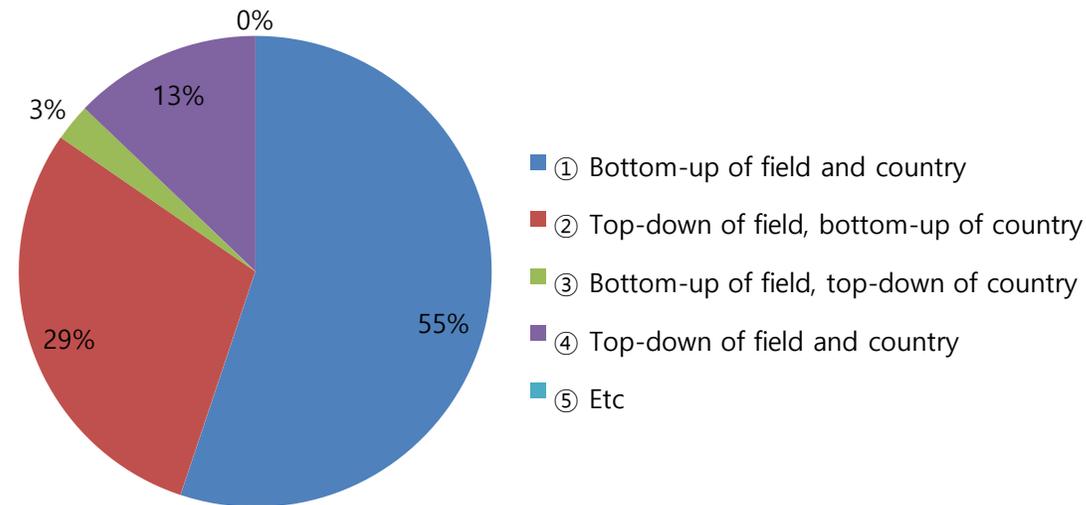
A Case: International joint research programs

Efficiency

- **The international co-performance has higher quality ?**

 - International joint research paper vs domestic research paper

<Excellence of international joint papers>

Classification	Total paper (MrnIF)	Joint research paper (MrnIF)	Domestic research paper (MrnIF)	Improvement of quality
Global Laboratory Program	78.39	81.85	74.38	10.04%
Global Research Network Program	69.66	73.75	64.36	14.59%
Energy International Joint Research Program	71.59	76.77	68.11	12.71%
Cf. Basic Research Program	71.68	75.48	70.21	7.51%

* Based on NTIS performance DB

Efficiency

- **The international co-performance has higher quality ?**

  - International joint patent (including foreign inventor or applicant) vs domestic patent

<Excellence of international joint papers>

Classification	Ratio of excellent patent (SMART)			
	Total patent	Joint patent	Domestic patent	Improvement of quality
Energy International Joint Research Program	20.0%	-	20.0%	-
Global Laboratory Program	18.9%	26.1%	17.5%	8.6%p
Global Research Network Program	17.6%	40.0%	15.2%	24.8%p
(Cf) Government R&D	12.5%	15.1%	12.4%	2.7%p
(Cf) Private R&D	11.9%	27.4%	11.6%	15.8%p

* Based on NTIS performance DB

Effectiveness

- **Do the programs achieve the objective?**

  - Effectiveness analysis is very difficult to calculate the improvement of research capacity

  - Here, the increase of quality of international joint performance means positive effect to the improvement of research capacity

❖ **Impact (long-term outcome) could be dealt in Effectiveness, but not due to the difficulties of measurement and its application to budget allocation.**

A Case: International joint research programs

Effectiveness

<Paper Citation Counts ('09-'14)>

Fields	Global average (No. of paper)	GLP (No. of paper)	GRNP (No. of paper)	EIJRP (No. of paper)
Biology&Biochemistry	9.61(397,159)	9.18(29)	6.23(61)	5(3)
Chemistry	8.33(881,860)	21.82(227)	12.26(59)	24.54(103)
Clinical Medicine	7.46(1,460,892)	8.26(21)	9.17(23)	-
Engineering	4.72(632,120)	11.70(18)	10.59(81)	4.43(48)
Environment/Ecology	7.665(237,240)	14.33(16)	3.33(3)	35.05(9)
Geosciences	6.987(231,392)	7.00(64)	15(1)	3.5(2)
Immunology	11.28(138,835)	10.03(15)	3.75(8)	-
Material Science	7.05(396,739)	16.17(120)	18.90(50)	7.5(77)
Molecular Biology&Genetics	14.62(240,922)	51.35(26)	13.67(26)	-
Multidisciplinary	26.55(14,958)	246.44(67)	7.17(16)	4.5(10)
Physics	7.135(656,729)	22.43(146)	7.89(134)	15.05(61)

* Based on fields publish more than 10 papers in Global Laboratory Program

A Case: International joint research programs

Systematic nature (management of programs)

<Volume of matching fund ('11-'15) (unit : million won>

Classification	GRNP	GLP	EJRP
Total Research Budget	27,120	103,806	77,611
Matching fund	-	58,873 (65.0%)	6,957 (9.0%)
Abroad Expenditure	8,418 (31.0%)	15,422 (14.9%)	17,835 (23.0%)

* Based on research proposal

Policy Direction of R&D Evaluation

- **3rd Basic Plan for R&D Performance Evaluation (2016~2020)**

  - Blueprint for R&D performance evaluation (confirmed by NSTC, 2015)

  - Encouraging quality and ripple effect based performance evaluation

  - Emphasizing on simplification of evaluation process and relieving evaluation burden

    for researcher oriented environment

Main Strategy	Sub strategy
Researcher centered evaluation	Inducement of creative & challenging research
	Simplification of evaluation process
Quality based evaluation	Enhancement of qualitatively excellence
	Focus on R&D program ripple effect analysis
Linkage btw policy-R&D investment-evaluation	Introduction of cluster evaluation (for a set of programs)
	Strengthening autonomy of line ministries in evaluation

Facing Issues for R&D Innovation in Korea

- How to evaluate individual R&D project/program to encourage creative and challenging research

- How to change evaluation governance : Balancing between **Autonomy & Accountability** (researcher Vs. ministry Vs. Coordinator)

- **The results of evaluation** has been directly used in budgeting, but **R&D evaluation should be more than that.**

- **Impact** (including ripple effect) aspect in the R&D evaluation should be put more emphasis on.

Impact Assessment in EU

- **Impact** can be defined as the effect of the project on its wider environment, and its contribution to **the wider policy or sector objectives** (as summarized in the project's overall objective).

- **Impact Assessments** examine whether there is a need for EU action and analyze **the possible impacts of available solutions**.
- These are carried out **during the preparation phase**, before the Commission finalizes a proposal for a new law.
- They provide **evidence** to inform and support the **decision-making process**.

Future R&D evaluation in Korea

- Focus on **Impact** in R&D evaluation can induce **Policy Evaluation.**

Performance

High Performance Low Impact	→	High Performance High Impact

Impact

Low Performance Low Impact		Low Performance High Impact
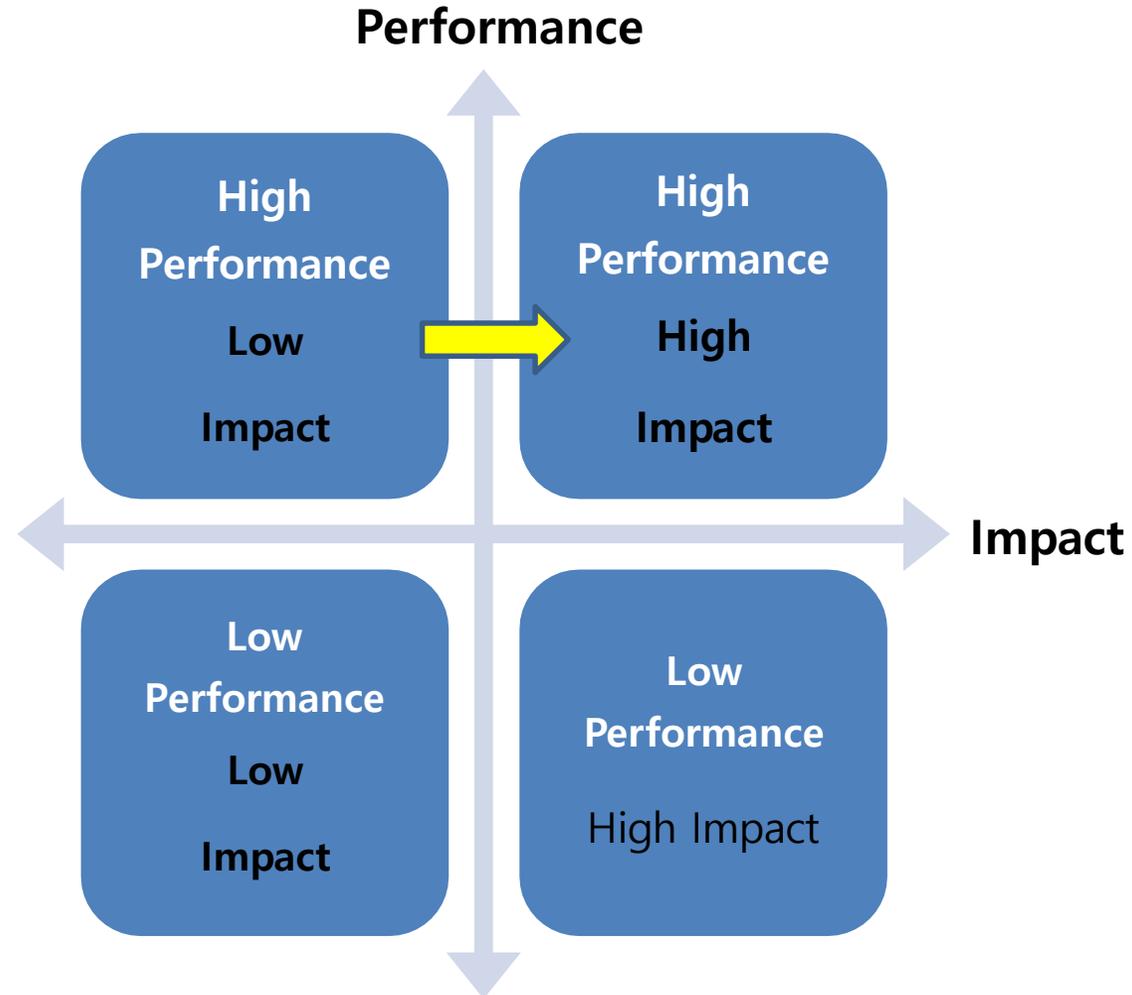
Thank you

jinwon529@kistep.re.kr
Jinwon529@k-erc.eu

The evolving research evaluation agenda

- The research evaluation agenda is shifting
  - Evaluation has historically been focussed on academic impact and 'excellence'
  - Excellence is important but selectivity alone can result in concentration that reduces research diversity
  - Stakeholder focus has shifted from research quality (academic impact) to research delivery (socio-economic impact)
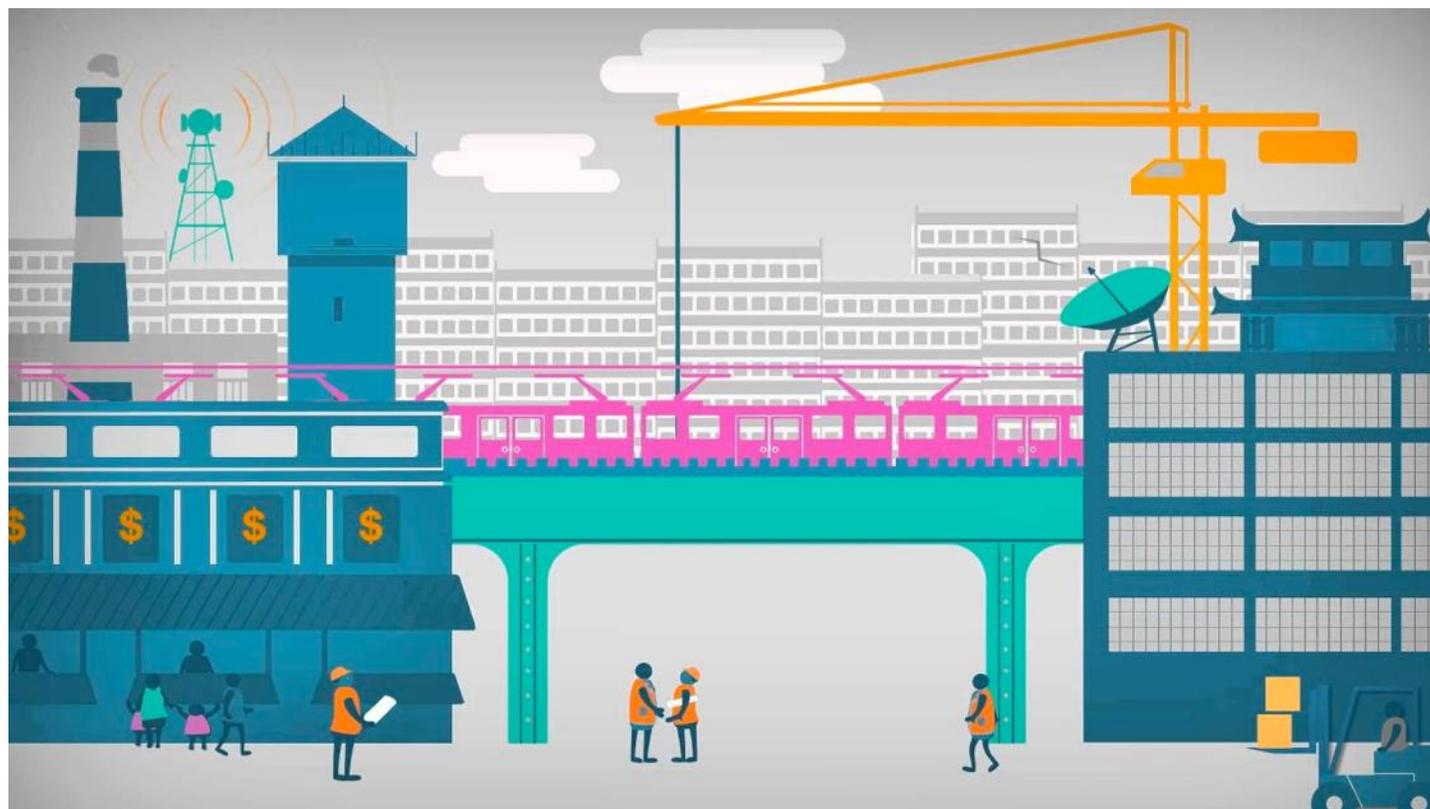
National portfolios have become more balanced

- The plot shows the unevenness of publications (according to WoS Category) for various nations over the last 40 years

- A lower Gini coefficient means the portfolio is more balanced (i.e. evenly distributed number of papers across categories)

- Most countries are converging around the same value and display a downward trend over time – except for the UK and US
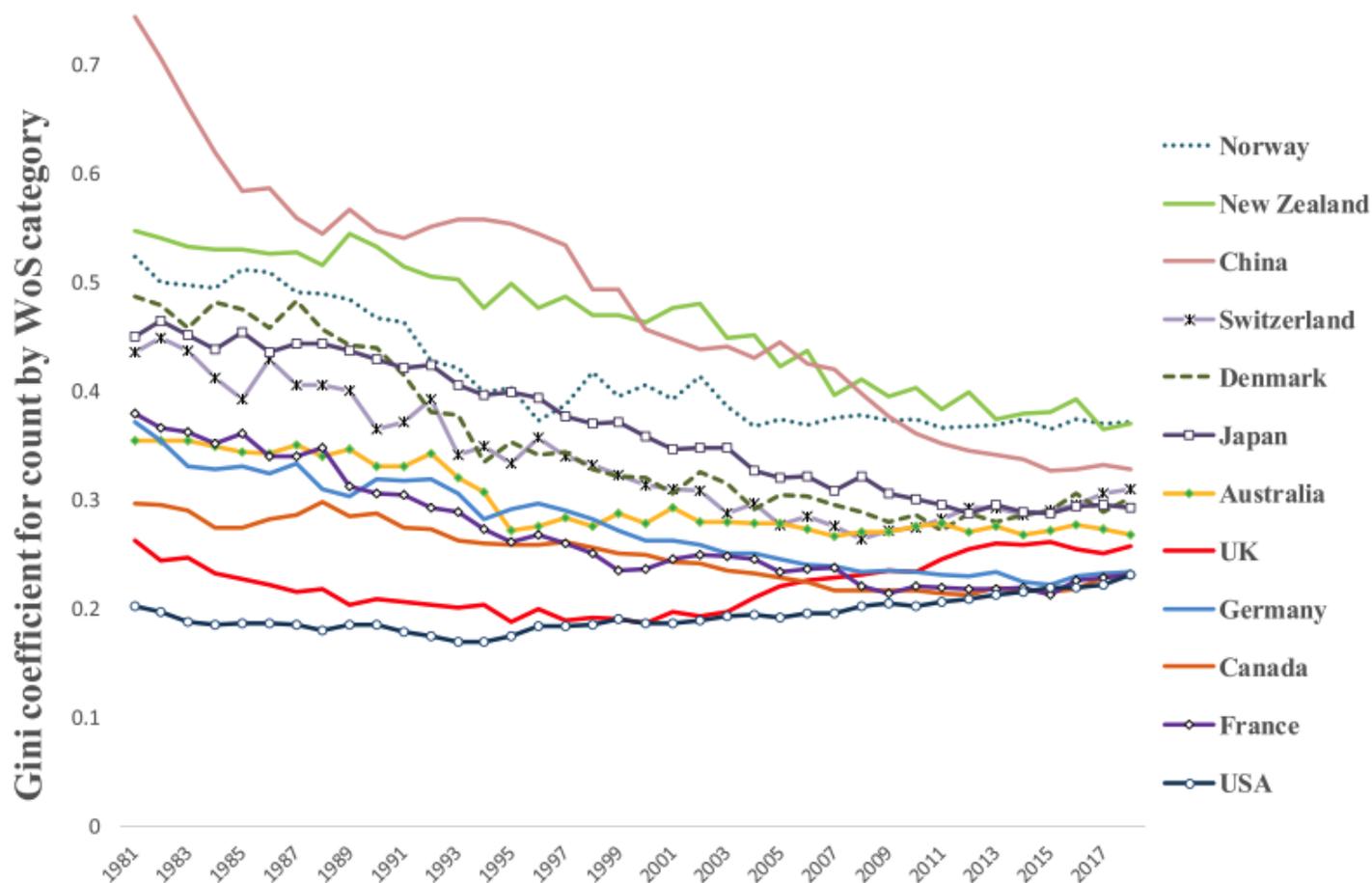
Figure 6. Gini coefficient values based on publication frequency data normalized against global baseline for the disciplinary diversity of papers with at least one national author for a country, calculated using the Web of Science (194 categories) data set.

Adams, J., Rogers, G., Smart, W., & Szomszor, M. (2020). Longitudinal variation in national research publication portfolios: steps required to index balance and evenness. Quantitative Science Studies, 1–38. https://doi.org/10.1162/qss_a_00073

History of Research Evaluation

UK Perspective

1990
- Advisory Board for the Research Councils
- UK science budget funding and output data

RAE1992
- Research Assessment Exercise

RAE1996
- How to make fair funding decisions?
- Benchmarking international research

RAE2001
- How to check submitted output is correct?
- Is selective funding too concentrated?

RAE2008
- Can metrics replace peer review?

REF2014
- Research Excellence Framework
- Introduction of case studies of socio-economic impact

REF2021
- Balanced approach to peer review and metrics
- Change to submission system

1991
- Work with ISI on National Science Indicators

1997
- Mapping and indexing UK research.

1998
- Adams J. Nature, 396, 615-618.

2001
- Validation of RASE database
- Fundamental review of selectivity and concentration
- Subject reviews
- Maintaining research diversity

2007-8
- Research assessment systems in UK universities
- Pilot project to test metrics across universities
- Strategically important subject review

2014
- Development of impact case study database
- Report on impact diversity with King's College, London

2018
- Data supply
- Advice to REF panels on correct use of metrics
- Verification of submitted outputs

Many countries are now including impact in the assessment process

- UK National Research Excellence Framework (REF)

  – REF2014 20% based on impact (25% in REF2021), reported via case studies

- Excellence in Research for Australia (ERA)

  – Introduced in 2018, similar model to UK but distinguishes engagement and impact

  – More fine-grained collection of impact and engagement types than REF

- European Commission

  – In H2020, periodic and final reports must state socio-economic impacts of the project

  – Open Access, Open Data and Open Science are import aspects to the research setup

- Canada

  – Embedded impact assessment, strong heritage in health and education research

- New Zealand

  – Subject-based, cyclical, similar indicators to REF and ERA

  – Policy language now strongly focussed on socio-economic and cultural impact

- Also developments in China, Japan, Singapore, South Korea, Netherlands, Ireland, and others

- US does not have centralised or block funding, hence national assessment has not evolved to include impact and outcomes

Global trend to assess research excellence AND impact

Complex array of options with various pathways to impact

Types of Impact	Political, Economic, Societal, Technological, Legal, Environmental, Health, Cultural …
Outcomes	Changes in behaviour and attitudes, health benefits, increased economic activity …
Beneficiary Groups	Students, Patients, Schools, Communities, Women, Policymakers, Citizens …
Evidence	Patents awarded, spin-out companies started, citations from grey literature documents such as clinical guidelines, testimonials, media coverage …
Reporting Mechanisms	Case studies, Funder reporting, Institutional collection (e.g. press-office, research-office, consultancy groups)

isi | Institute for Scientific Information

Case Study Model
Examples from UK (REF) and Australia (ERA)

https://impact.ref.ac.uk/casestudies/

https://dataportal.arc.gov.au/EI/Web/Impact/ImpactStudies

Headline findings from analysis of REF2014 case studies

1. The societal impact of research from UK HEIs is considerable, diverse and fascinating
2. The research underpinning impact is cross-disciplinary, and the benefit arising from research has multiple impacts
3. UK HEIs have a global impact
4. The quantitative evidence supporting claims for impact is diverse but *inconsistent*, suggesting that the development of robust impact metrics is unlikely
5. The impact case studies provide a rich resource for analysis, but the information was built (by researchers) for *assessment* purposes and may need to be restructured for *analysis* purposes
6. The *interpretation* of impact will continue to evolve
7. Socio-economic impact is no more certain or predictable than other research outcomes

The nature, scale and beneficiaries of research impact

An initial analysis of Research Excellence Framework (REF) 2014 impact case studies

King's College London and Digital Science

March 2015

Grant, Jonathan, and S. Hinrichs. "The nature, scale and beneficiaries of research impact: An initial analysis of Research Excellence Framework (REF) 2014 impact case studies." *HEFCE-Higher Education Funding Council for England* (2015).

isi | Institute for Scientific Information

Impact Topics

- Topic modelling used to extract salient concepts in the body of the impact case study text

- Chords connect co-occurring topics with width proportional to the number of case studies that reported them

isi | Institute for Scientific Information
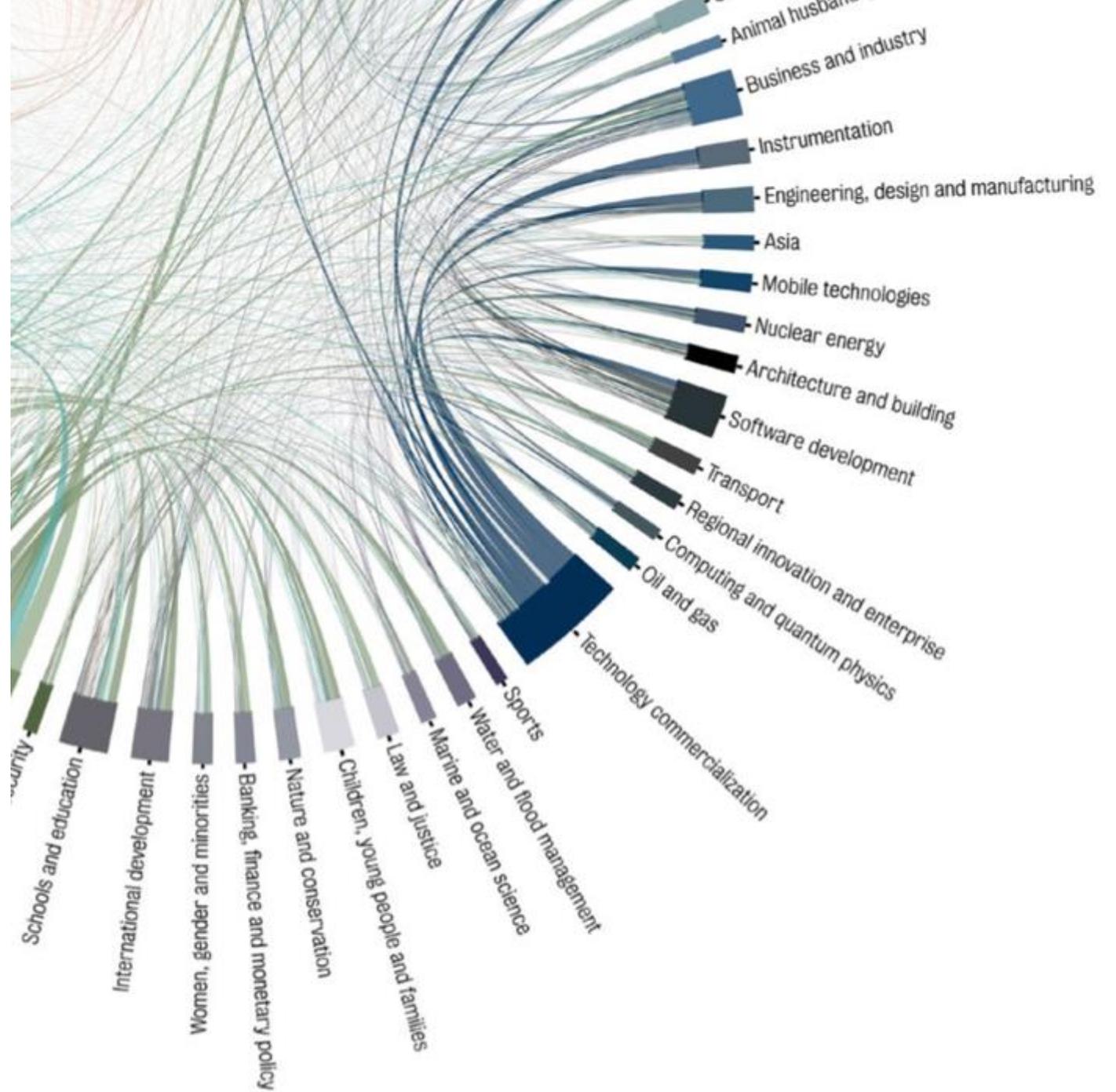
Beneficiaries of the impact
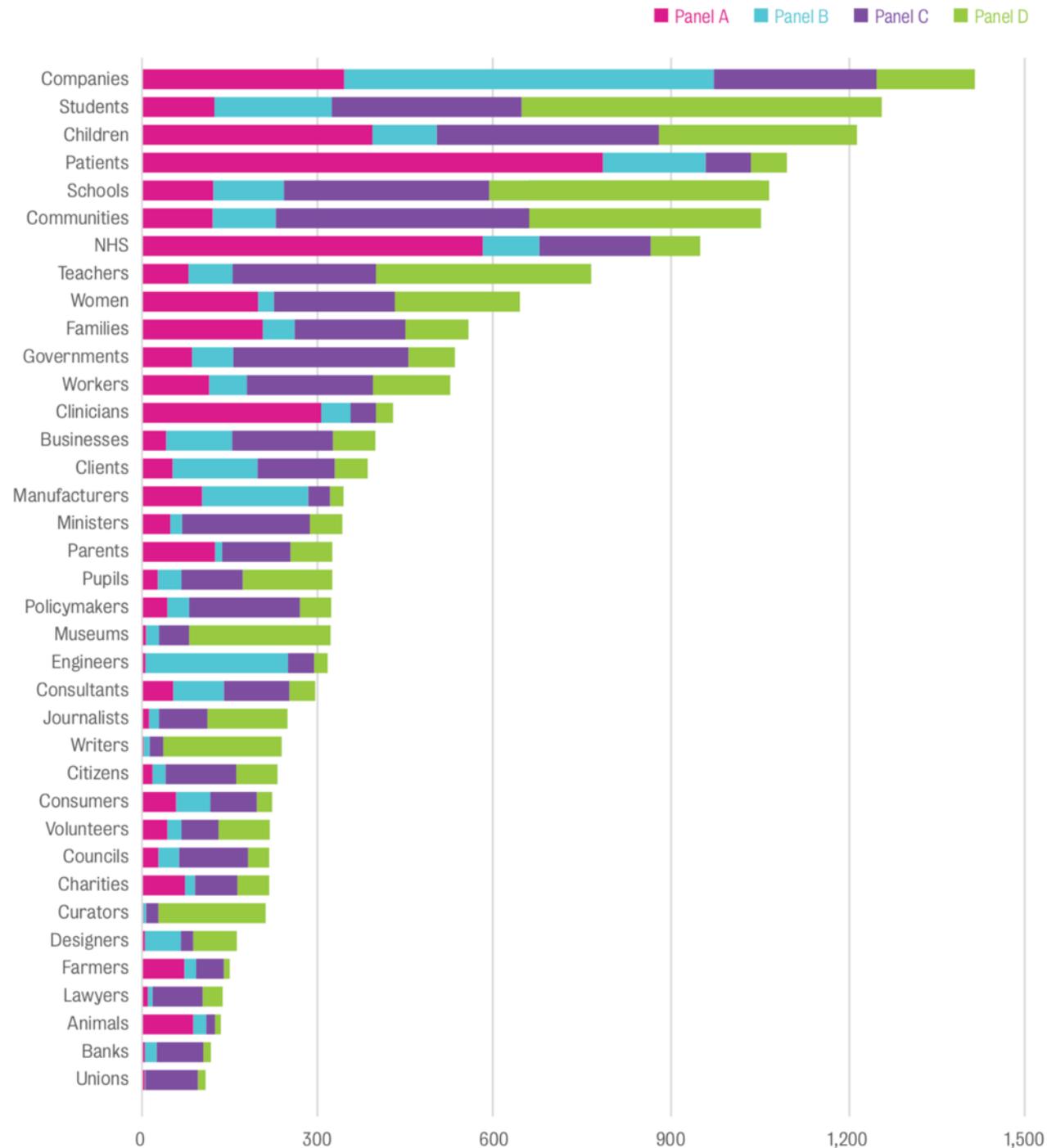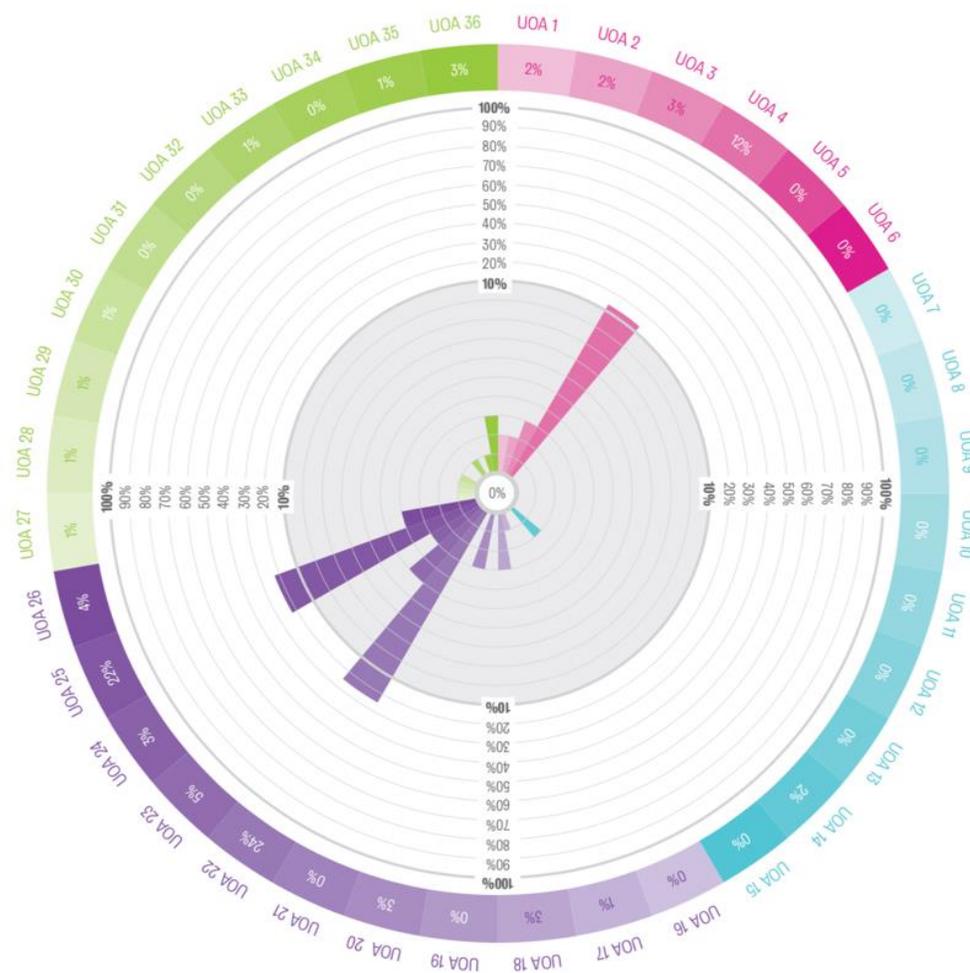
('Children, young people and families', n=198)

# A closer look at evidence

Text mining impact case studies

- The Impact Case Studies describe the research that underpins the impact reported via references to articles and grants awarded
- A range of other evidence is also quoted in the final section (Sources to corroborate impact)
- We used text mining to categorise these additional references to understand the range of evidence used and investigated their correlation with final scores
- A broad typology was developed (see table) through systematic analysis of evidence cited using search patterns (e.g. regular expressions, url matching)

Type	Description
Testimonial	A letter or email from an individual or organisation describing the utility of the research
Report	Any grey-literature reference such as a policy document, white paper, parliamentary proceeding, etc.
Article	News articles (print or online), blog posts and other forms of web content
Media	TV / Radio appearances, online videos, podcasts, etc.
Activity	A workshop, conference, exhibition, social event, etc
IP	Intellectual property and trademarks
Award	Honours and other forms of recognition
Legal	Legal proceedings

isi Institute for Scientific Information

Top evidence types included in REF2014 Impact Case Studies

According to Main Panel

- Testimonials are the most widely used type of evidence followed closely by Reports

- The Arts and Humanities disciplines relied much more on Media (TV / radio, online videos, podcasts, etc) and Activities (social events, exhibitions, workshops) than the other 3 panels

- Patents were mentioned in 632 / 6637 case studies, mostly in Panels A + B

- A diverse range of grey literature (Report) was mentioned across the panels
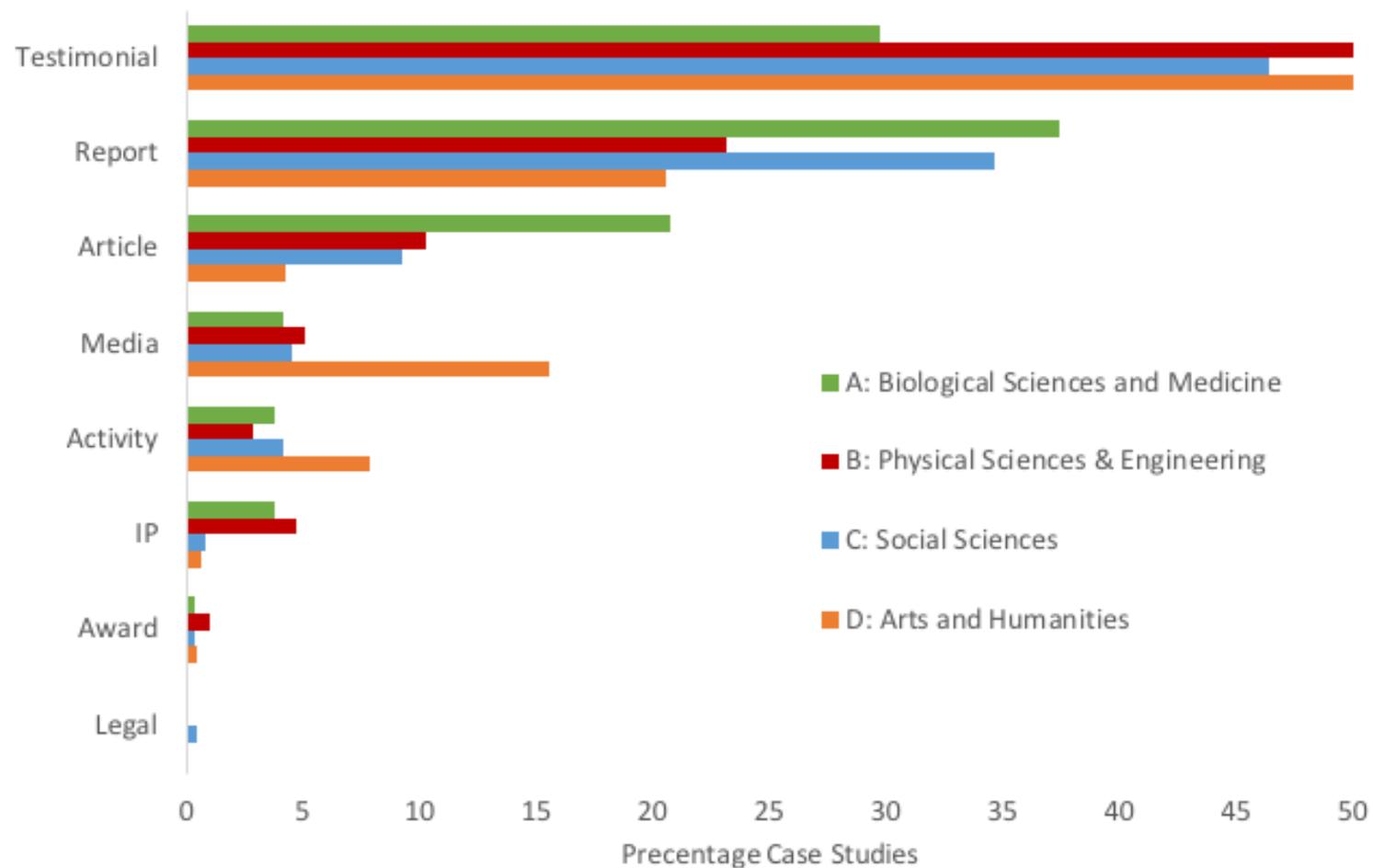
The correlation between evidence types and scores

- Scores are only available at the Unit of Assessment & Institution level only

- This means that we know the percentage of 4*,3*,2*,1* case studies that a university submitted to a particular panel, but we can't say what any particular case study scored

- At a high level, we are able to assess the association between scores and evidence types

- The correlations between the amount of evidence of a given type, and the Grade Point Average (GPA) score for a set of case studies is shown in the coloured matrix

- A value of 1 implies maximal positive correlation (spearmen), 0 no correlation, and -1 a maximal inverse correlation

	A: Biological Sciences & Medicine	B: Physical Sciences & Engineering	C: Social Sciences	D: Arts & Humanities
Testimonial	**-0.15**	**0.04**	**0.08**	**0.17**
Report	**0.19**	**0.11**	**0.15**	**0.08**
Article	**0.19**	**0.09**	0.02	-0.01
Media	-0.01	**0.07**	**-0.07**	0
Activity	-0.03	-0.02	-0.04	**-0.06**
IP	0.05	0.05	-0.01	0
Award	**-0.06**	0.01	0.01	0
Legal	-0.03	0	0	0

The values in bold are significant (p value < 0.05, where the null hypothesis is that the indicative score and the amount of a given evidence type are uncorrelated)

isi | Institute for Scientific Information

The correlation between evidence types and scores

- Panel A shows a positive correlation between score and use of reports as evidence: it might be expected that much health impact is corroborated by policy documents and clinical guidelines

- Panel B & C scores are also positively correlated with the use of reports as they show evidence of social, economic and environmental policy outcomes

- This contrasts with panel D, where testimonials are the evidence type most positively associated with score

- In fact, testimonial count is negatively correlated with score for Panel A

- For Panel B, scores are positively correlated with a range of evidence types, but the coefficients are small

- Articles (mostly news coverage) are equally correlated with score as reports in Panel A

	A: Biological Sciences & Medicine	B: Physical Sciences & Engineering	C: Social Sciences	D: Arts & Humanities
Testimonial	**-0.15**	**0.04**	**0.08**	**0.17**
Report	**0.19**	**0.11**	**0.15**	**0.08**
Article	**0.19**	**0.09**	0.02	-0.01
Media	-0.01	**0.07**	**-0.07**	0
Activity	-0.03	-0.02	-0.04	**-0.06**
IP	0.05	0.05	-0.01	0
Award	**-0.06**	0.01	0.01	0
Legal	-0.03	0	0	0

The values in bold are significant (p value < 0.05, where the null hypothesis is that the indicative score and the amount of a given evidence type are uncorrelated)

Extending our view of the information system

Research is linked to outcomes via a range of evidence types
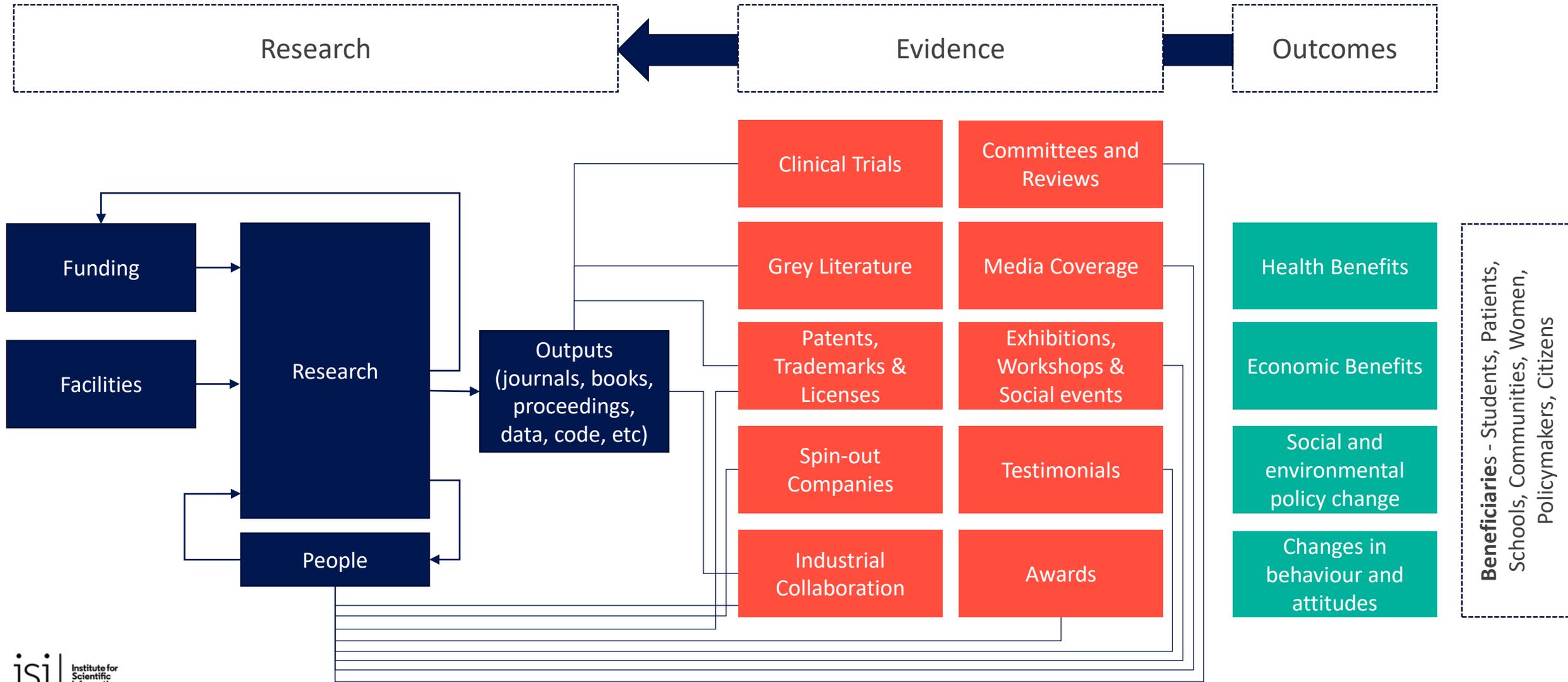
Other Resources

The **Gateway to Research (GtR)** website has been developed by UK Research and Innovation (UKRI) to enable users to search and analyse information about publicly funded research.

https://gtr.ukri.org/

Research Outcomes are collected via ResearchFish and associated with individual grants awarded by UK Research Councils

This an excellent resource to understand how certain types of research lead to particular outcomes

The taxonomy is developing and provides clues on how the typology of impact is emerging

The outcomes tree can be expanded to show more details

The Future of Research Evaluation

- Increased focus on research impact in national assessment programmes and funder initiatives

  - Peer review will remain a crucial component, supplemented by bibliometrics

- More structured capture of engagement types

  - As we discover more about the evidence types used, cataloguing and tracking systems can be improved. Free-text capture requires extensive data-mining for post-evaluation analysis

- Development of ontologies for capturing impact

  - These will be domain specific, already quite advanced in medicine

  - Opportunity to mobilise academic societies and professional bodies to establish their own view

- Increased researcher awareness of the need to track impact
- Improved platforms for researchers to evidence various engagement and impact activities

  - Funding awarded, editorial and peer-review duties, speaking engagements, news and media coverage, awards, consultancy activity, industrial collaboration, grey-literature citations

- Better understanding of how academic research is used outside of established scholarly channels such as grey literature citations

Virtual tour through Krakow
&
Lounges

(12.45-14.00)