Indicators in a science policy context

Peter van den Besselaar

Science System Assessment, Rathenau Instituut
&
Vrije Universiteit Amsterdam

Science policy – the issues

- Main issues
  - Forecasting
  - Evaluation
  - Funding
  - Governance and organization
  - Careers
- Increasing demand for evidence based policy
- Which requires
  - models and theories for understanding and derived indicators for informing policy
  - As older indicators are not always useful: data based instead of theory based

Science System Assessment

- Science and science policy studies for evidence based science policy
- Broad program on topics mentioned
- Output: relevant evidence (or indicators) for policy makers; in variety of forms
- Collaboration with stakeholders:
  - Ministry of Education Culture and Science
  - Research Council (NWO)
  - Universities and research institutes
  - Royal Academy and other advisory bodies
  - Other stakeholders

Indicators in policy context

- Different types of indicators – different types of use
  - Indicators for informing policy
    - Comparison of research institutions, countries, etc.
  - Indicators for evaluation
    - Research evaluation
    - Indicators in grant allocation processes
    - Indicators for strategic positioning
  - Indicators for foresight
    - Indicators for strategic positioning
    - Mapping knowledge dynamics
- Two examples of studies we did
  - Not only answers, also many remaining questions
  - Indicators for science policy as a research domain

Part 1:
Indicators for grant allocation

Policy background

- Policy question (MaGW): do we select the best researchers for our funding?
- Policy assumption: selecting the best researchers is possible and useful to do
  - Is this valid?

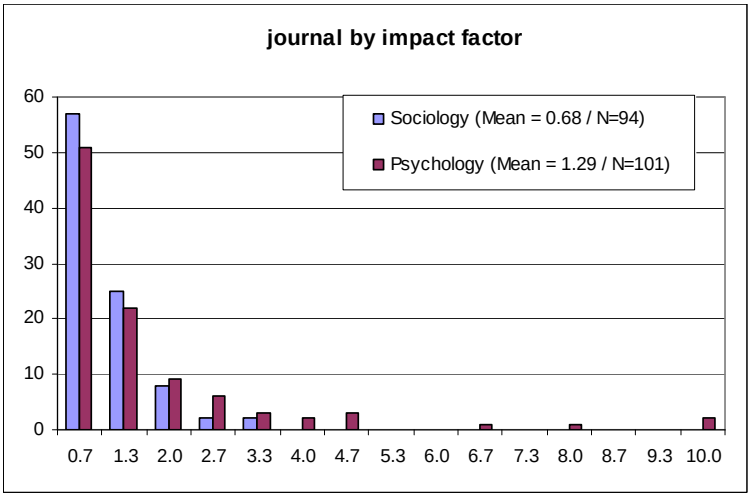
Indicators for research performance in this study

- Productivity
  - Publications in ISI indexed journals
- Visibility of work
  - Received citations (3 years window)
    - and what about other indicators: keynotes, editorial boards, awards?
- What about other output?
  - Non-journal output: e.g., books
  - Non-English output
  - Doctorates
  - Networks
  - Societal impact (public sphere, consulting, policy)
- However:
  - the definitions of the actors involved (here the council) was leading
  - After obtaining the results, the discussion emerges (probably leading to new definitions and related indicators!

Social Science Citation Index

- Always problematic to use, but also always used!
- The usual issues
  - Biased coverage for the social sciences
  - English language bias
  - Differences in citation behavior by subfield (see next slide)
  - Mistakes (spelling!)
    - Estimated: in at least 10% of the records
  - However: We do not evaluate individuals, but the evaluation of individuals by the allocation system (individual mistakes do not matter)

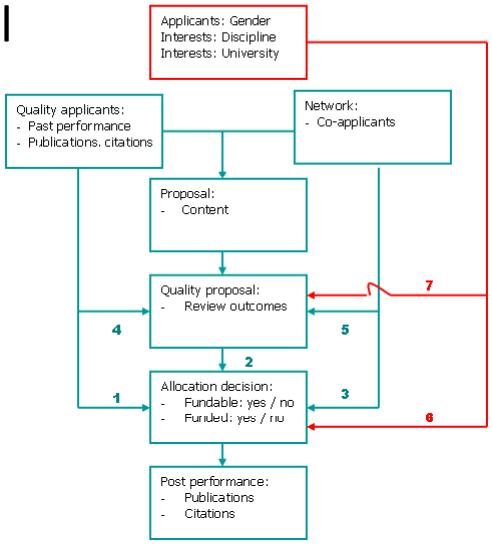
Sociology & psychology: different behavior



Approach / model

- Merton:
  - peer review based on scientific norms → best researchers
    - past performance (of applicant and co-applicants) and peer judgment determine allocation process
- Latour
  - peer review based on interests/power → powerful researchers
    - Status research field and university; gender; networks relations with other important actors

Model



From data to indicators

- WoS indicators
  - Publications: All SSCI records (A, N, L, R) with an address in the Netherlands: 2001-2005.
  - Total citations received to these publications – per Feb 9 2007
  - More than 10.000 records, these were matched with the 1290 MaGW records
  - Matching through last name and first initial
  - Of course: some error
- Peer review data
  - All review results – also preliminary evaluations
  - Recoded to a similar 5 point scale
- Result: per applicant a two-dimansional 'quality indicator'
  - his/her past performance in terms of publications and citations
  - His/her average referee score (and number of referees)

Methods

- Relation between the performance indicators and the decision of the Council established through:
  - Visual inspection
  - Analysis of variance
  - Correlation analysis
  - Discriminant analysis

Successful vs. unsuccessful applicants (1)

Sign = .000		N	Mean	St Dev	95% Conf Interval	
					Low	Up
Cit	Success	275	36	71	28	44
	Failed	911	16	45	13	19
	Total	1186	20	53	17	23
Pub	Success	275	4.4	5.99	3.7	5.2
	Failed	911	2.7	4.92	2.4	3.0
	Total	1186	3.1	5.23	2.8	3.4

Successful vs. unsuccessful (2)

Sign = .000		N	Mean Referee score	St Dev	95% Conf Interval	
					Low	Up
	Success	274	1.59	0.63	1.52	1.67
	Failed	904	2.68	1.05	2.61	2.75
	Total	1178	2.43	1.06	2.36	2.49

Funding, citations, publications and peer review: correlation?

		Cit	Ref	euro
Pub	Pearson Correlation	.818**	-.173**	.159**
	Sig. (2-tailed)	.000	.000	.000
	N	1186	1178	1186
Cit	Pearson Correlation		-.179**	.197**
	Sig. (2-tailed)		.000	.000
	N		1178	1186
Ref	Pearson Correlation			-.326**
	Sig. (2-tailed)			.000
	N			1178

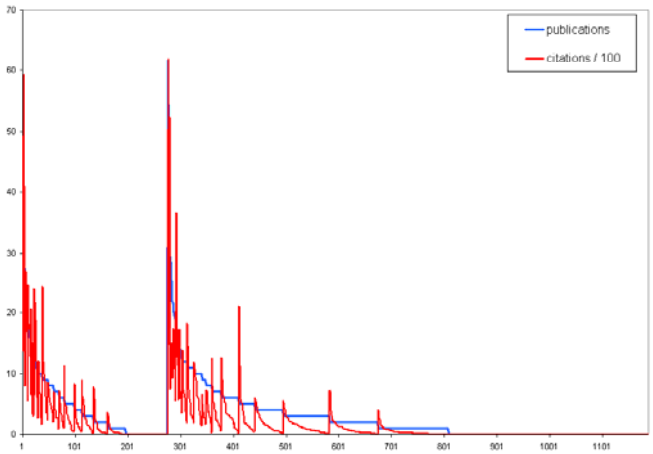
Can we predict success by past performance and peer review?

		Success versus Failed	Predicted Group Membership		Total
			Success	Failed	
Original	Count	Success	231	43	274
		Failed	354	550	904
	%	Success	84.9	15.7	100.0
		Failed	39.2	60.8	100.0

Stepwise, *Cit* and *Ref* in analysis

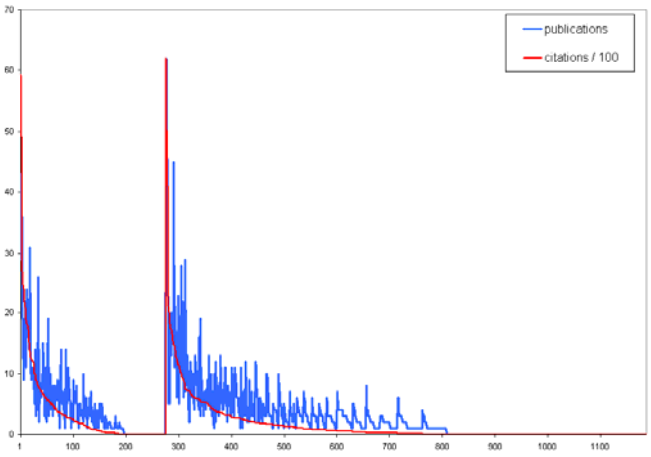
31-8-2009 Peter van den Besselaar / Summerschool 17

Performance and success 2003-2005 (1)

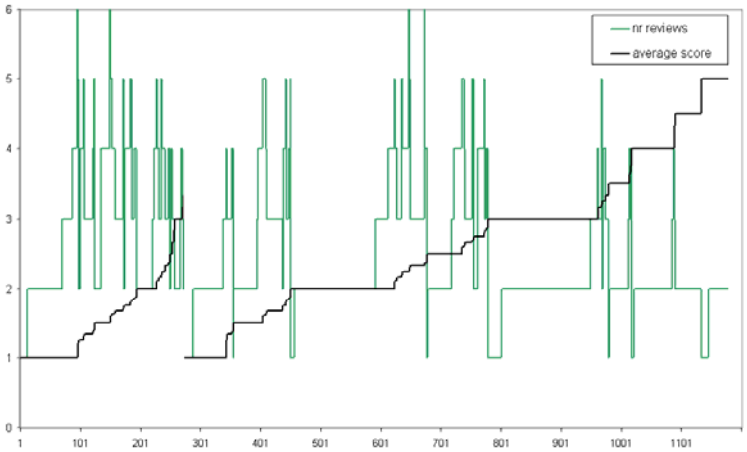


31-8-2009 Peter van den Besselaar / Summerschool 18

Performance and success 2003-2005 (2)



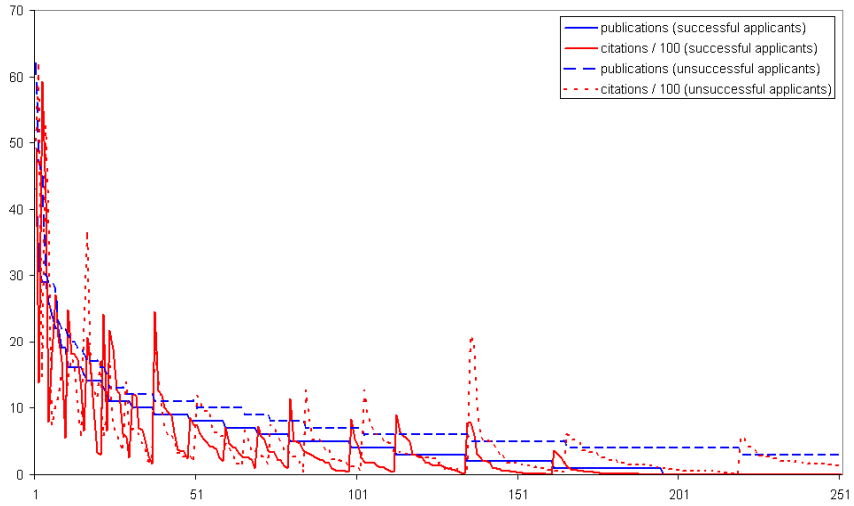
Peer review and success



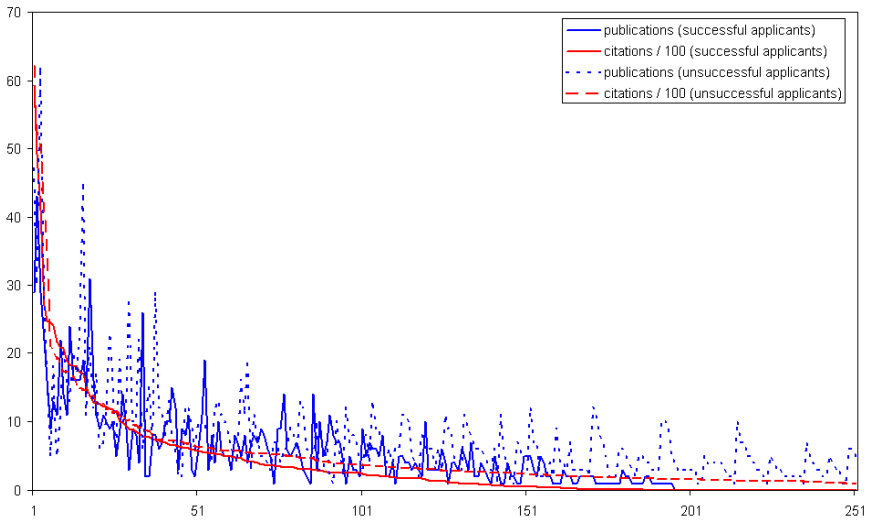
And only the top?

- 275 granted applications
- 275 best rejected (in terms of past performance?

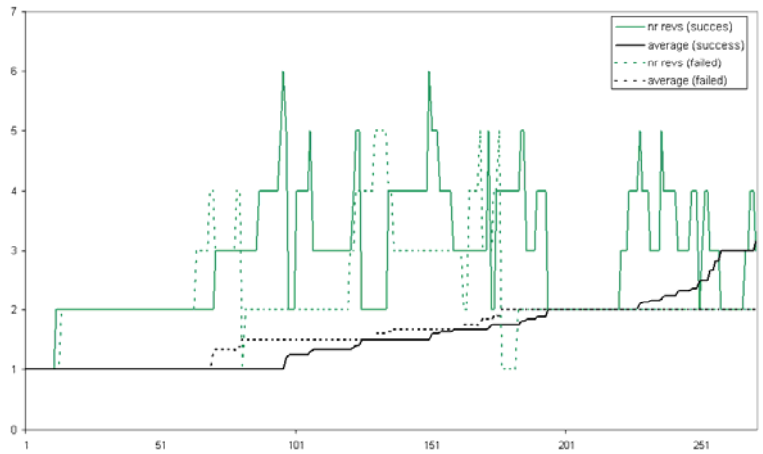
Publications: top rejected win



Citations: top rejected win



Referee scores? Top rejected equal



Successful vs. unsuccessful TOP applicants: opposite effect

		N	Mean	St Dev	95% Conf Interval	
Cit	Success	275	36			
Sign = .000	Failed	277	48			
	Total	552	42			
Pub	Success	275	4.4			
Sign = .000	Failed	277	6.9			
	Total	552	5.7			
Ref	Success	276	1.62			
	Failed	271	1.64			
	Total	547	1.63			

31-8-2009 Peter van den Besselaar / Summerschool 25

Successful vs. unsuccessful TOP applicants (2): opposite effect

552		Cit	Ref	Euro
Pub	Spearman's rho	.833**	-	-.256**
	Sig. (2-tailed)	.000	ns	.000
Cit	Spearman's rho		-	-.262**
	Sig. (2-tailed)		ns	.000
ref	Spearman's rho			-
	Sig. (2-tailed)			ns

31-8-2009 Peter van den Besselaar / Summerschool 26

Can we predict the winners in the top?

- NO!

Some tests

- Does differentiation in sub-disciplines makes a difference? HARDLY
- Does differentiation between instruments make a difference? HARDLY (and in an unexpected direction)
- Network: co-applicants and their past performance
  - Slight increase with better co-applicants

Findings

- Whole set: correlations are
  - weak between PP and funding
  - moderate between Rev and funding
  - low between Rev and PP:
    - quality is ambiguous concept
    - And indicators should reflect this, as should the use of indicators
- Top half
  - at best no correlations (or negative)
- Confirmed by other studies!

substantial conclusions:

- Do the grants go to the best researchers? No!
  - Defined in the way we did:, the answer is: to the good researchers, not necessary the best. Removing the weak tail words reasonable.
- Could the procedure be improved to select the best? No!
  - Correlation between past performance and review scores is very low: identifying the best seems an impossible task. Criteria remain fuzzy and subject to interpretation.
  - Picking the winners on the individual level therefore seems not feasible - and a too rationalistic model of funding decisions.
- Is that a problem? No!
  - Funding should improve knowledge production as much as possible
  - No reason to expect the highest marginal increase of return from funding the best researchers (if we are able to identify them). Evidence even suggests the opposite!
  - Criteria for a good funding system: allowing for much variation (creativity, innovation); unbiased selection * negative selection (removing) of the lower part

Implications for indicators

- Bibliometric indicators do not correlate with each other
- Bibliometric indicators do not correlate with decision making:
- Probably not useful FOR EVALUATION and POLICY at INDIVIDUAL level.

- More focus on evaluating institutions like funding than on evaluating individuals/groups
  - Does the system support good and useful research?

- New indicators needed for evaluating effectiveness of research funding schemes – at systems level:
  - Measures for adequate variation
  - Measures for adequate selection

- Much discussion provoked by the study!

PART 2

MAPPING KNOWLEDGE DYNAMICS FOR FORESIGHT

The need for decisions

- Science policy in an heterogeneous environment: different visions, positions, interests, values.
- Main issue: distribution of resources over research fields and over types of research (curiosity driven, basic, mission oriented, applied) in this heterogeneous environment.
- New funding instrument: large investments in big mission oriented programs.
- Decisions are based on claims and promises about opportunities and risks of existing and new scientific developments.
- Foresight needed?

Foresight in a (science) policy context

- Royal Academy foresight committee's:
  - Where are research fields moving? What should be the future agenda? Implication for agenda setting and funding
- Departmental 'knowledge chambers':
  - Matching future knowledge needs with research agenda's
- Rathenau Instituut:
  - Foresight of risks of new science and technology
  - Identifying promising developments in new and existing research fields

Science System Assessment: research program

- Long term aim: understanding knowledge dynamics
  - models and typologies on different levels
  - Indicators for types of knowledge dynamics -> recognizing
  - methods and tools for data collection, handling, integration, analysis, and visualization.
- Short term agenda: 'foresight studies'
  - applying models and tools in case studies
  - to inform science policy

Understanding the dynamics of research fields in context

- Knowledge dynamics: three indicators
  - Growth
  - Position in the landscape: (new) disciplines, interdisciplinary fields, multidisciplinary developments
  - Relations in the landscape: knowledge flows between fields (cognitive dependencies)
- Contexts of knowledge dynamics
  - Institutions, funding instruments, research evaluation
  - Policy, agenda setting
  - Economic and societal demand

Why not only actors / experts?

- Foresight cannot any more depend on experts only:
  - Science system has become very large
  - Even experts lack overview, and have often different interpretations
  - Experts have also interests
  - Risk of groupthink
- But: Intensive interaction with experts and other stakeholders. Expert meetings, focus groups
  - information source
  - validating the results
  - framing the conclusions

Methodological issues of indicators for knowledge dynamics

- Delineation of research fields
- Fuzzy boundaries
- Change in research fields and change of research fields

- Journals or papers (or even other document types: books, proceedings, electronic documents)
- Citation patterns, co-citation patterns, co-word patterns
- Indicators for multi/interdisciplinary knowledge integration
- Indicators for transdisciplinary collaboration between research and practice (for agenda setting and knowledge dissemination and use)

Example 1: cognitive neuroscience

- Increasing political discussion about the state of the educational system. No consensus.
  - Cognitive neuroscience claims new developments useful for solving these problems: e.g., brain development and curriculum structure
  - Research funding in the NL: investments in thematic areas with scientific potential and social relevance.
  - Foresight study about cognitive neuroscience to provide an answer.

(1) Femke Merkx, Reinoud van Kooten, Thomas Gurney, Peter van den Besselaar, The development of transdisciplinary learning sciences. Rathenau Instituut 2009.

Example 2: coastal research

- Policy debates about coastal defense in relation to climate change and ecological preservation.
  - Changing ways of framing the problem: defense against water (seawalls) versus accommodating and integral coastal zone management: new policy challenges.
  - How does the research system reacts? Do funding systems invoke changes in agenda's?
  - Foresight study about coastal defense to provide an answer (1).

(1) Femke Merkx, Peter van den Besselaar, Positioning indicators for cross-disciplinary challenges. *Research Evaluation* **17** (2008)

Example 3: development of new and existing fields

- What is the state of a new research field such as regenerative medicine (2), or an existing field such as communication studies (1)?
  - What are the hot topics?
  - How is it moving internationally?
  - And what is the position of local research institutes?

(1) Peter van den Besselaar, A map of information, media and communication studies. Rathenau Instituut 2007

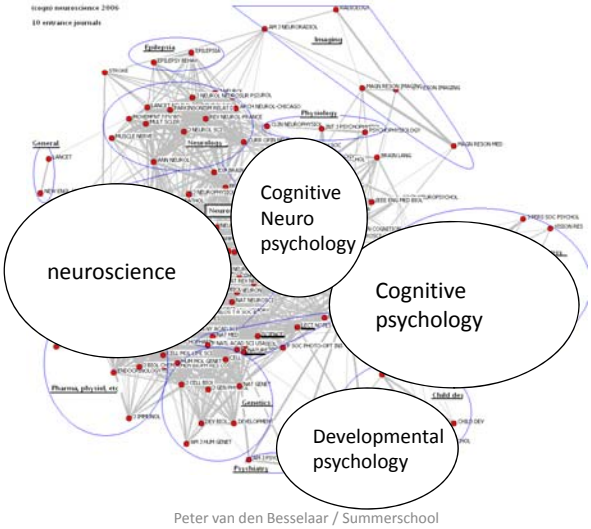
(2) Peter van den Besselaar, Thomas Gurney, Regenerative medicine as an emerging field. Rathenau Instituut 2009

Indicators for the integration of educational sciences & brain research

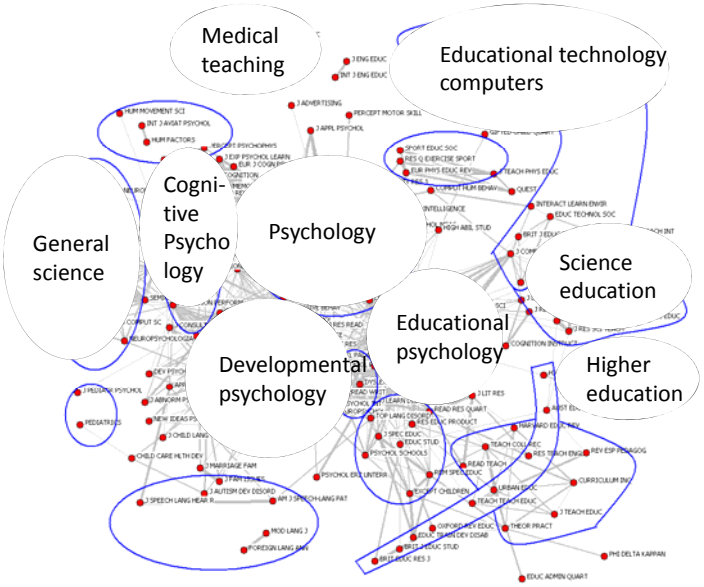
- Understand the dynamics of knowledge in context, may help to inform decision making in specific cases.

- Are cognitive neurosciences and the educational sciences at some spots converging?
- Is a multidisciplinary research field of cognitive educational science emerging?
- Does knowledge flows between the educational sciences and cognitive/brain sciences?
- Are contextual factors favorable (eg collaboration)?

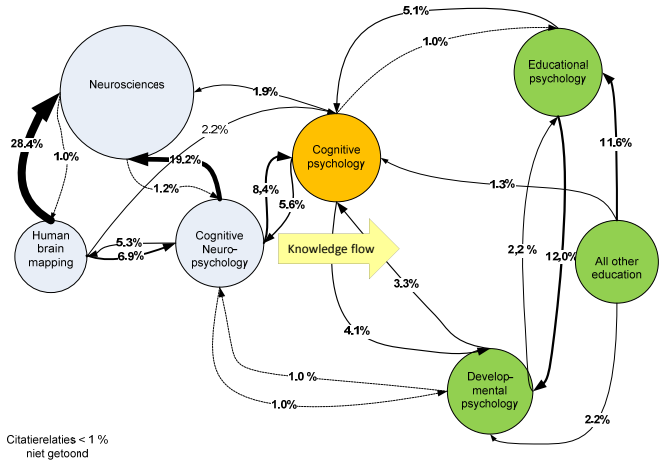
Neuro/brainscience: education invisible



Education research: brain invisible



Knowledge flow is marginal



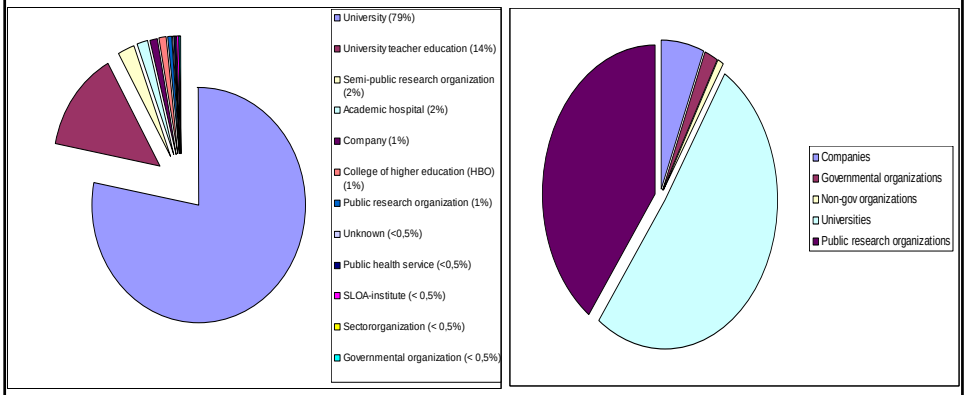
Findings

- Brain and neuroscience and education sciences:
  - No convergence
  - No interaction
- Policy implications: forms of funding, agenda setting and organizing the early promises?

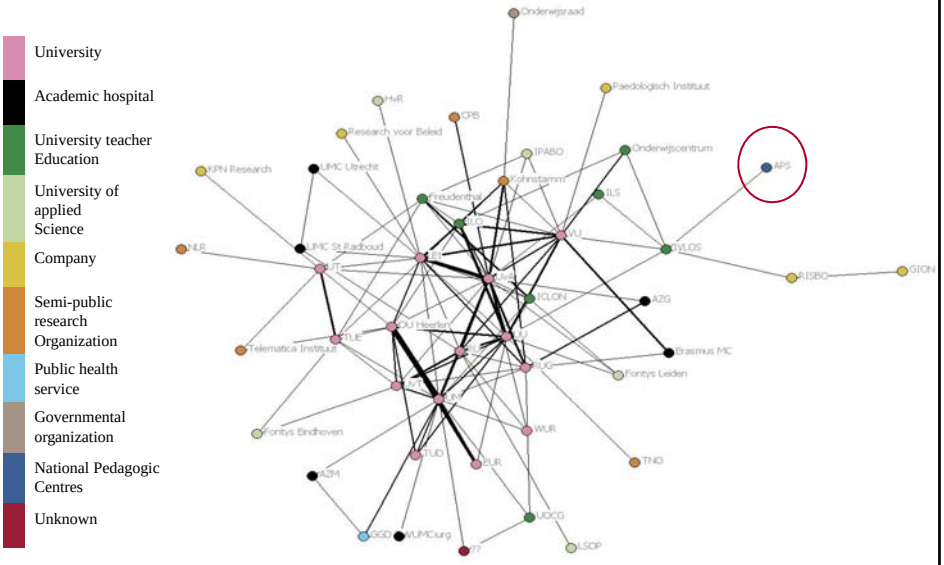
Coastal research gives a different picture:

– A few multidisciplinary journals integrating
- 'hard coastal engineering', 'soft coastal engineering', coastal management and policy.
- Social sciences are taking part, but on a small scale.

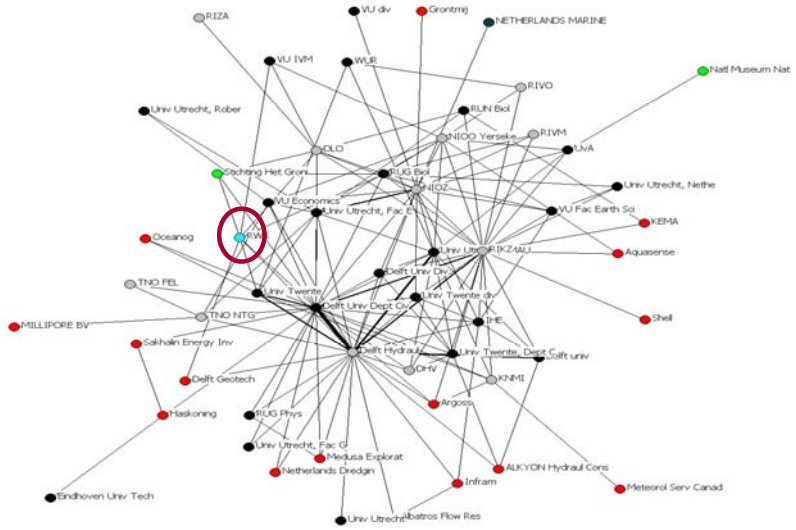
Collaboration is different in education research and in coastal research



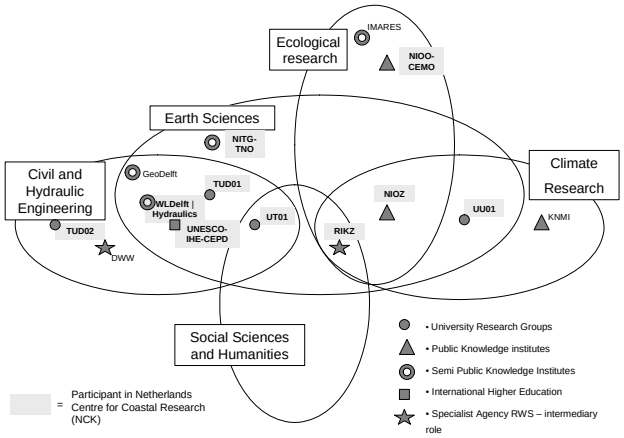
Co-authors network: education



Co-authors network: coastal research



Transdisciplinary & interdisciplinary collaboration in coastal research



Lessons second case

- We are developing useful indicators for:
  - Developments of research fields
  - For knowledge integration
  - For interdisciplinary collaboration
  - For transdisciplinary (social and cognitive) networks
- Mapping of the knowledge dynamics informs foresight activities of stakeholders in supportive and in critical sense: identifying real developments and identifying hypes.
- Stage of development informs policies: funding, organizing, agenda setting, evaluation – how is still only partly clear. So indicators need theoretical underpinning