

JULY 5



INTERNATIONAL WORKSHOP

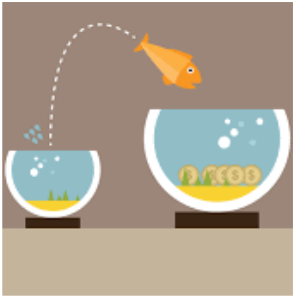
"Technology Upgrading and Catching Up in Latin America: Reflections from Policy Evaluations"

STI Policy Evaluation in Latin America: evidences based on Science and Innovation Policy Evaluation Repository (SIPER)

Adriana Bin

Laboratory for Studies on the Organization of Research &
Innovation (GEOPI) / School of Applied Sciences / University
of Campinas





STI Policies for Technology Upgrading and Catching Up

(tech and non tech capabilities, sustainability, inclusive change, governance, GVC, organizational change...)

Effective policy making depends on evidence, learning and planning



Evaluation contributes to determine the effects of policy interventions (and also to understand intervention rationale and implementation)

How STI Policy evaluation is being conducted and what is its real role in policy formulation and implementation?





STI Policies for Technology Upgrading and Catching Up

Effective policy making depends on evidence, learning and planning



Evaluation contributes to determine the effects of policy interventions (and also to understand intervention rationale and implementation)

How STI Policy evaluation is being conducted and what is its real role in policy formulation and implementation?



META-EVALUATION is “*the evaluation of evaluation*” (Scriven, 1969)

“the process of delineating, obtaining, and applying descriptive information and judgmental information—about the utility, feasibility, propriety, and accuracy of an evaluation and its systematic nature, competent conduct, integrity/honesty, respectfulness, and social responsibility—to guide the evaluation and/or report its strengths and weaknesses” (Stufflebeam, 2001)



STI Policies for Technology Upgrading and Catching Up

Effective policy making depends on evidence, learning and planning



Evaluation contributes to determine the effects of policy interventions (and also to understand intervention rationale and implementation)

How STI Policy evaluation is being conducted and what is its real role in policy formulation and implementation?



META-ANALYSIS is “a form of quantitative synthesis of studies that address a common research question” (Stufflebeam, 2001)

“allows an improved comparison and understanding of interventions and their effects by taking into account the results of a large number of evaluations. (...) Provided that raw-data of evaluations is made accessible, the information given in a large number of such evaluations can be used as data input for subsequent analysis.” (Edler et al., 2008)

EVALUATION SYNTHESIS as “a modified form of meta-analysis (...) to assess the overall combined effects, redundancies, contradictions and remaining bottlenecks of policy interventions” (Edler et al., 2008)



STI Policies for Technology Upgrading and Catching Up

Effective policy making depends on evidence, learning and planning



Our **PURPOSE** is to present the results of an on-going initiative – Science and Innovation Policy Evaluation Repository (SIPER) concerning Latin American countries.



Evaluation contributes to determine the effects of policy interventions (and also to understand intervention rationale and implementation)

SIPER is a meta-evaluation exercise and it can contribute to meta-analysis and evaluation synthesis exercises.

How STI Policy evaluation is being conducted and what is its real role in policy formulation and implementation?



Few comprehensive meta-evaluation exercises of STI Policies (Edler et al., 2012; Ancaiani et al. (2015), Felix et al. (2017), Boyce (2017) focusing on HEI; Weißhuhn et al. (2017), Weißhuhn et al. (2017) and Knudsen (2018) focusing on agricultural and health research; Heijs & Vergara (2011) and Porta (2014) focusing in LA.

Outline

1. What is science and innovation policy evaluation repository – SIPER and SIPER-LA
2. What are the findings (so far)
3. What can we say about STI Policy Evaluation practice in LA
4. What is ahead (research agenda)

**WHAT IS SCIENCE AND INNOVATION
POLICY EVALUATION REPOSITORY
(SIPER)?**

About Science and Innovation Policy Evaluations Repository (SIPER)



RISIS

Research infrastructure for research and innovation policy studies

This project is funded by the European Union under Horizon2020 Research and Innovation Programme Grant Agreement n°824091



The current datasets cover five critical dimensions: ERA dynamics (3 datasets), firm innovation dynamics (3 datasets), public sector research (3 datasets), research careers (3 datasets) and a **repository on research and innovation policy evaluations**.



SIPER

Science and Innovation Policy Evaluations Repository

About Science and Innovation Policy Evaluations Repository (SIPER)

SIPER is a central source of knowledge on science and innovation policy evaluations.

Its **aim** is twofold: (i) to **provide on-line access** to a unique collection of policy evaluations, located at a single location; (ii) **to allow policy learning** by providing an informed analysis of the database contents that is both searchable by policy makers and other stakeholders.

Our primary **audiences** are: **policy makers** concerned with the design, implementation, management and evaluation of science and innovation policy interventions and instruments; **evaluation practitioners**; and those engaged in **research** (in both the public and private sectors) into science and innovation policy evaluation and the broader area of Science, Technology and Innovation (STI) studies.

Dataset owner: **Fraunhofer Institute for Systems and Innovation Research** <https://www.isi.fraunhofer.de/en.html>
(originally developed by Manchester Institute of Innovation Research - MIIoIR)

Dataset manager: **Jakob Edler** jakob.edler@isi.fraunhofer.de

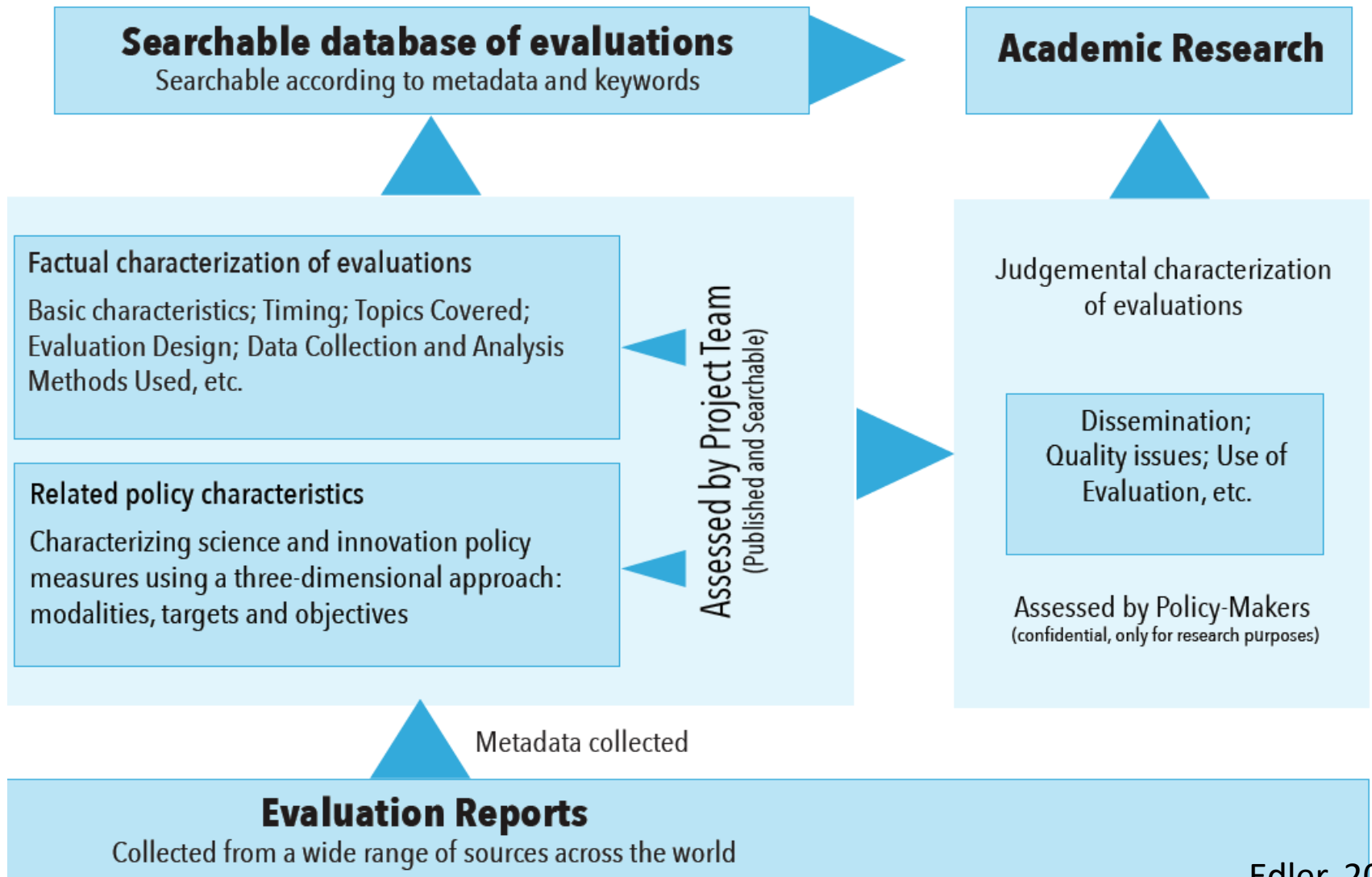
Brazilian team (covering LA): **University of Campinas (UNICAMP)** - Adriana Bin (coord.), Rafaela Andrade, Lígia Vasconcellos, Sergio Salles-Filho, Ana Maria Carneiro, João Pedro Rodrigues



About SIPER Evaluations

- Criteria applied to ensure they are appropriate for application of characterisation process:
 - Relate directly to a STI policy instrument (i.e. all public interventions that support STI activities)
 - Evaluate a specific instrument or group of instruments
 - Have a distinguishable methodology
 - Contain an analysis/judgement about the subject of the evaluation
 - Provide some sort of evidence of performance
- Inclusion/non-inclusion is not judgement of ‘quality’ of report or the evaluation/measure to which it refers

How SIPER is constructed



Policy Measure Characterization (PL)

- PL0 About the Policy Measure being evaluated
- PL1 Targets (Recipient of the support)
- PL2 Modalities (How support is provided)
- PL3 Explicit policy objectives (Why the support is provided)

Factual Characterization (FC)

- Section 0: About the evaluation
- Section 1: Basic information (conduction, timing, purpose, reference to intervention rationale)
- Section 2: Topics covered (aspect of the program that the evaluation examine)
- Section 3: Evaluation Design
- Section 4: Data Collection Methods
- Section 5: Data Analysis Methods
- Section 6: Quality Issues

Science and Innovation Policy Evaluation Repository (SIPER)

[Home](#) [Repository](#) [About](#) [Publications](#) [News & events](#) [Data & Methods](#)

[Contact](#)

SEARCH SIPER DATABASE >

What is SIPER?

SIPER is a unique central source of data and knowledge on research and innovation policy evaluations. Here you will find a comprehensive collection of evaluations on science and innovation policy instruments. Not only that, each policy measure and each evaluation in the SIPER database has been systematically characterised following a uniform template. On this basis, users can search for evaluations that show a certain combination of characteristics and/or deal with a certain kind of policy instrument or measure.

What is the scope?

We aim to include evaluations of science and innovation policy programmes conducted since 2000 (with a few exceptions). Starting with OECD countries, SIPER aims at a global geographical coverage.

Search in the database

www.si-per.eu

Select your search criteria

Hide









Please click on the search options panels to show/hide details of search criteria.

1: Related policy measure characteristics

2: Evaluation characteristics: Basic

3: Evaluation characteristics: Topics covered

4: Evaluation characteristics: Design

- 1 Experimental 
- 2 Quasi-experimental 
 - 1 Before/after comparison 
 - 2 Comparison/control groups 
 - 3 Beneficiary self-reporting 
- 3 Non-Experimental 
- 4 Included explicit comparison/benchmarking with similar measures 
- 5 Benchmarked against previous phases/evaluations of programme/measure 

5: Evaluation characteristics: Data collection methods

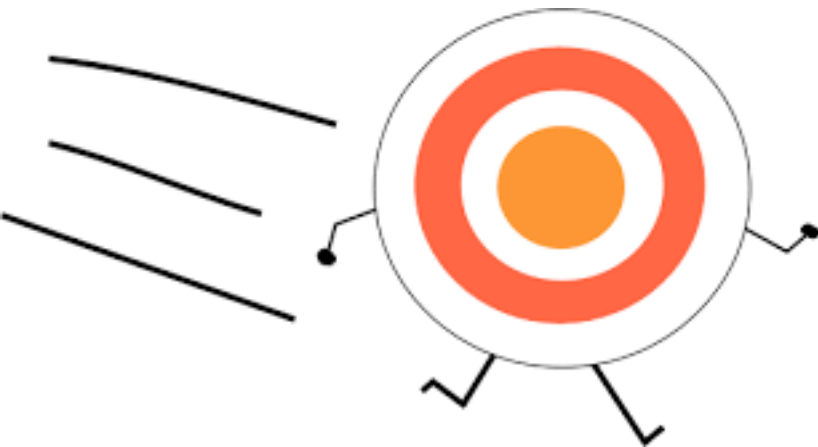
6: Evaluation characteristics: Data analysis methods

7: Document properties

Updates of SIPER database

SIPER database is in a process of **continuous updating**, with new evaluations being added.

Until the database contains a fully comprehensive set of evaluations, **any differences in numbers of evaluations per country will not reflect the actual number produced by that country** and is only an artefact of the search process.



Updates of SIPER database



TOTAL = 708

WORLD (- LA) = 565
(mainly Europe)

LATIN AMERICA = 143

Updates of SIPER - LA database



ARGENTINA = 26(+1)



BRAZIL = 37



CHILE = 27(+1)



COLOMBIA = 16



MEXICO = 19



URUGUAY = 17

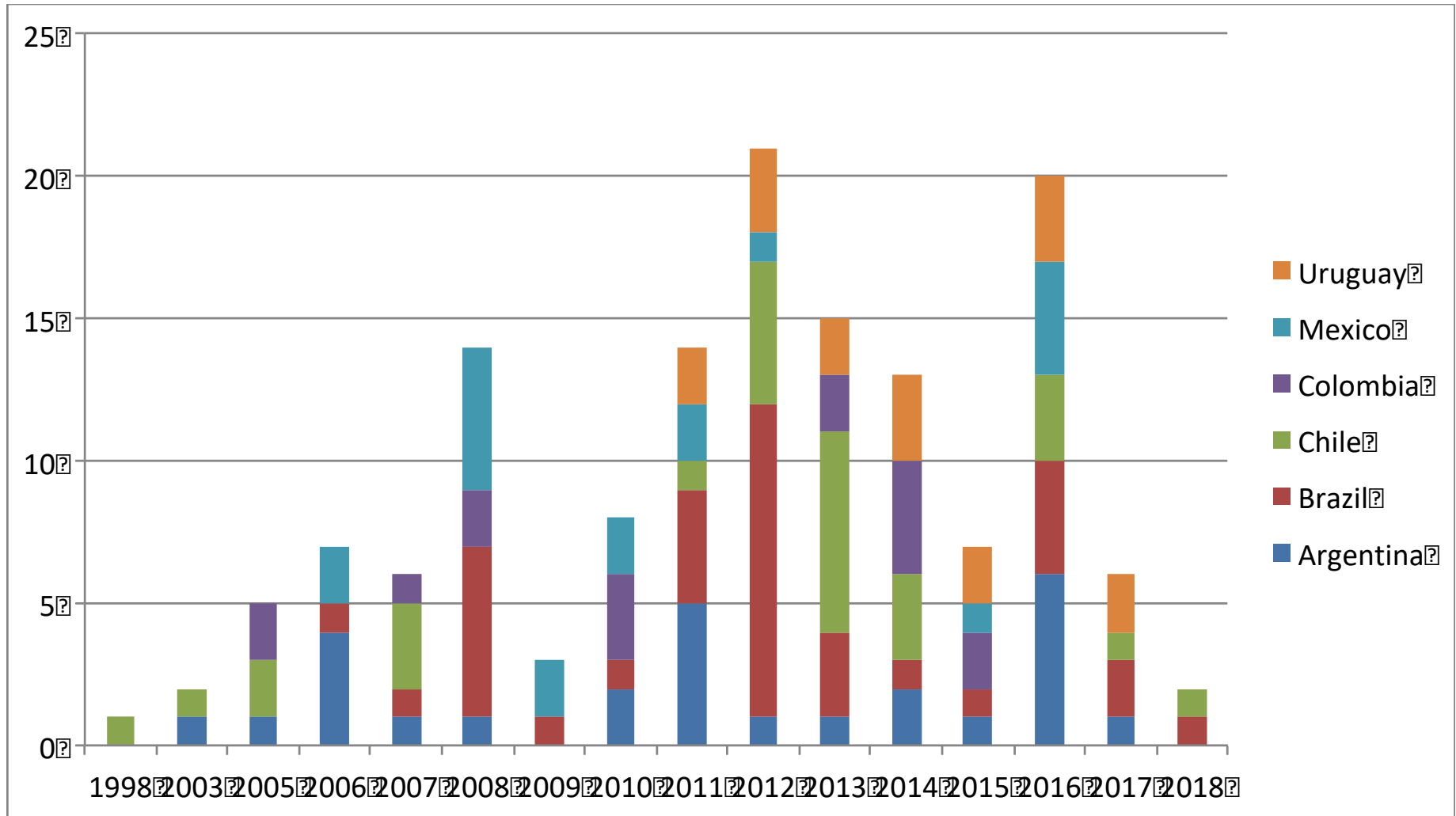
143 ALREADY IN
SIPER

*(but just 18 retrieved
through website search
– database in transition
from MloIR to
Fraunhofer)*

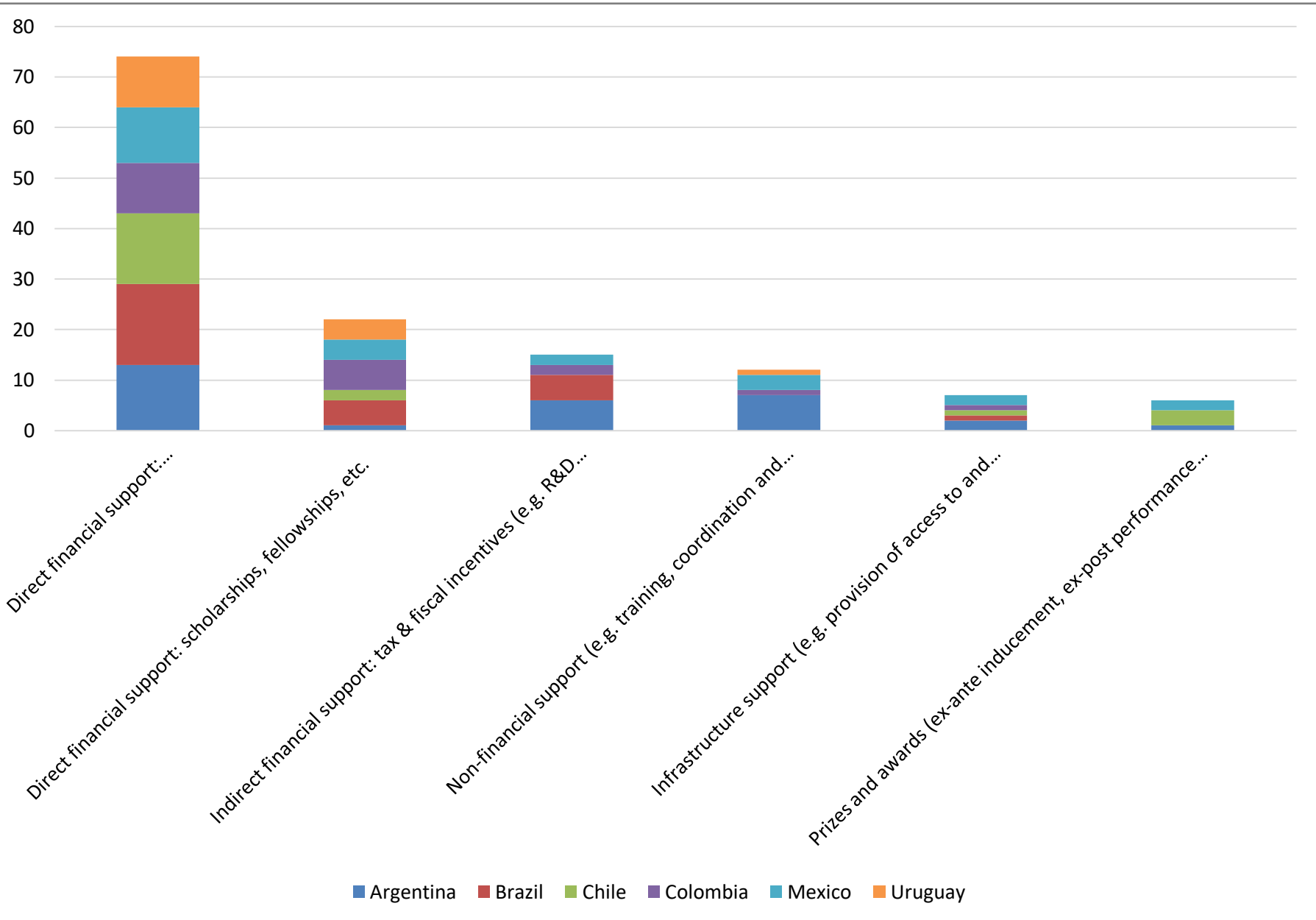
SOME MORE TO BE
UPLOADED AND
CHARACTERIZED...

SOME MORE TO BE
COLLECTED (from
these 6 countries and
others from LA...)

Updates of SIPER - LA database

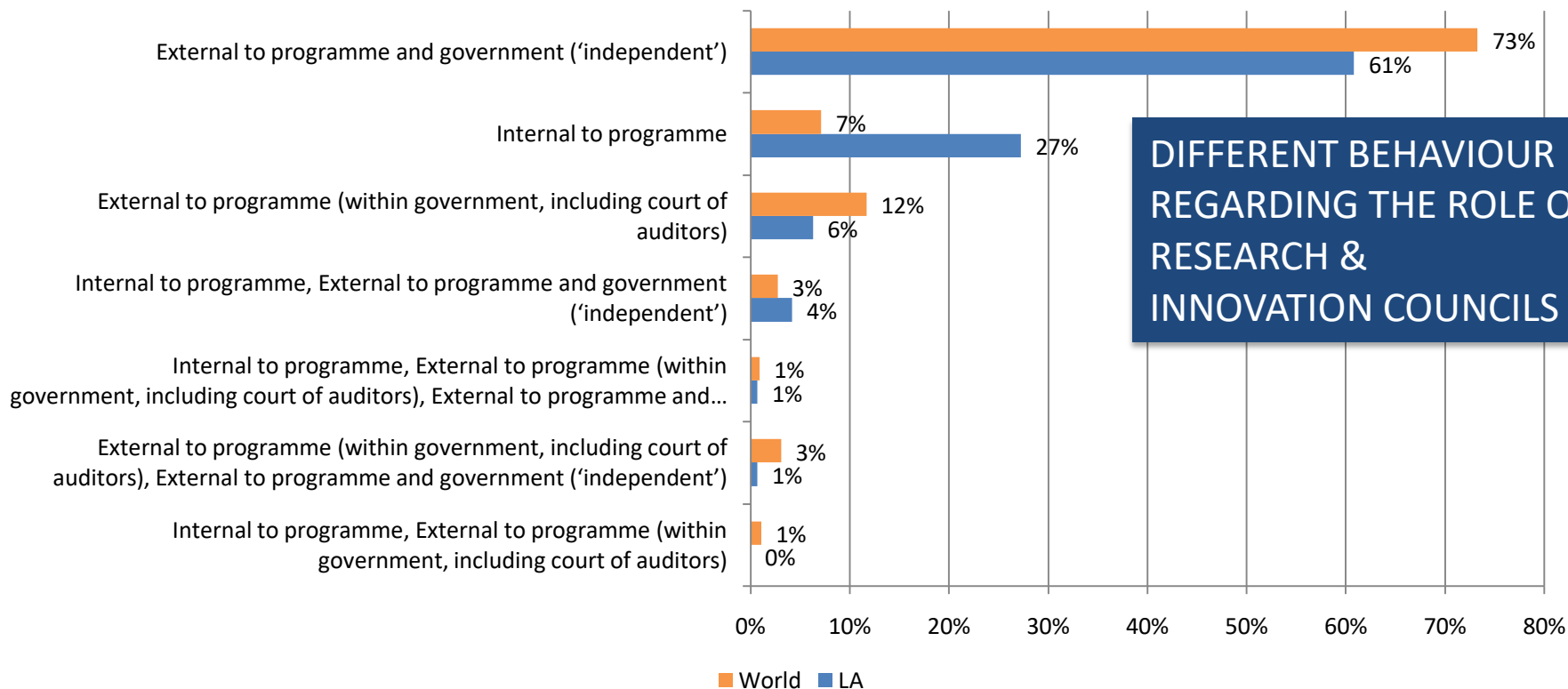


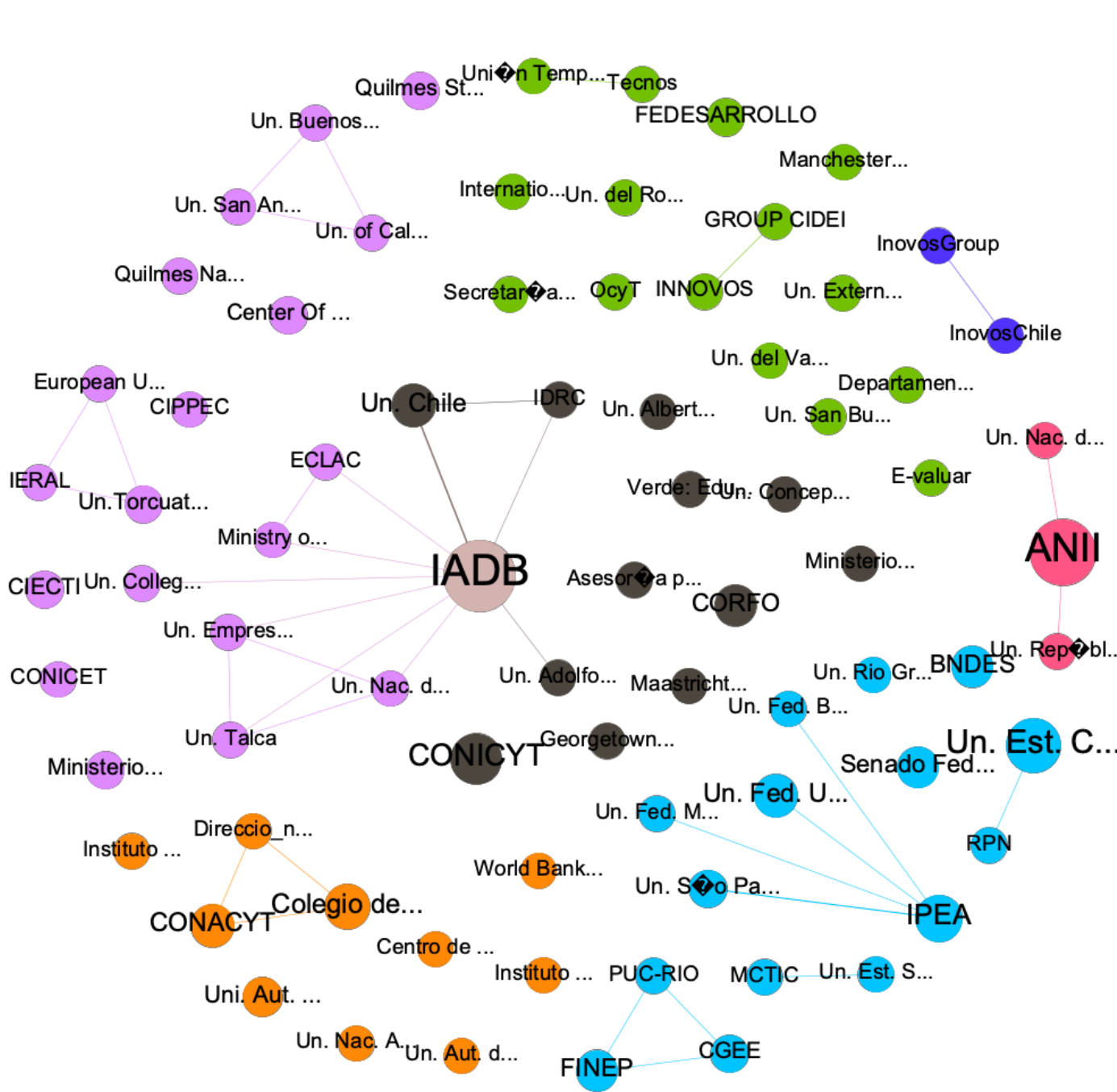
WHAT IS BEING EVALUATED?



**WHO CONDUCTED THE
EVALUATIONS?**

Who conducted the evaluation?	LA	World
Internal to programme, External to programme (within government, including court of auditors)	0	6
External to programme (within government, including court of auditors), External to programme and government ('independent')	1	17
Internal to programme, External to programme (within government, including court of auditors), External to programme and government ('independent')	1	5
Internal to programme, External to programme and government ('independent')	6	15
External to programme (within government, including court of auditors)	9	64
Internal to programme	39	39
External to programme and government ('independent')	87	401
	143	547





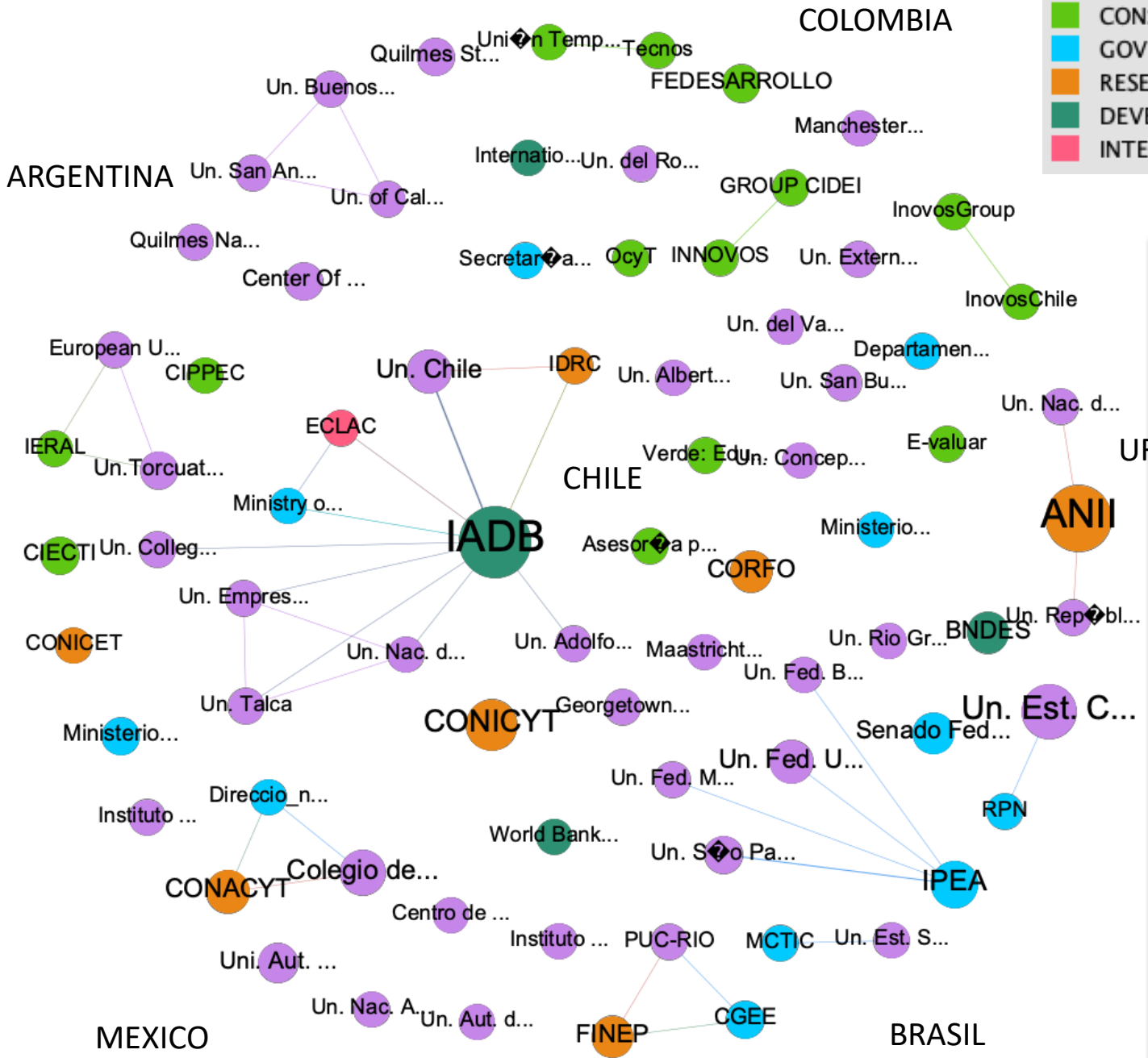
- Argentina
- Colombia
- Brazil
- Chile
- Mexico
- Uruguay
- Chile, Argentina
- Multiple Countries

THERE IS NO ESTABLISHED LATIN AMERICAN NETWORK OF EVALUATORS

IADB PLAYS A CENTRAL ROLE AS AFFILIATION INSTITUTION OF AUTHORS OF ALL COUNTRIES (EXCEPT FROM URUGUAY)

THERE ARE SOME NETWORKS INSIDE THE COUNTRIES (PARTICULARLY IN ARGENTINA AND BRAZIL)

- UNIVERSITY
- CONSULTING
- GOVERNMENT
- RESEARCH & INNOVATION COUNCIL
- DEVELOPMENT BANK
- INTERNATIONAL ORGANISATION



UNIVERSITIES AND CONSULTING (INCLUDING PRIVATE FOR PROFIT AND NON PROFIT) ARE THE MAJORITY – “EXTERNAL TO PROGRAMME AND GOVERNMENT”

RESEARCH AND INNOVATION COUNCILS AND DEVELOPMENT BANKS HAVE IMPORTANT ROLES, MAINLY IN URUGUAY, BUT ALSO IN ARGENTINA, CHILE AND MEXICO – “INTERNAL TO PROGRAMME”

**TIMING, PURPOSE AND REFERENCE TO
INTERVENTION RATIONALE OF
EVALUATIONS?**

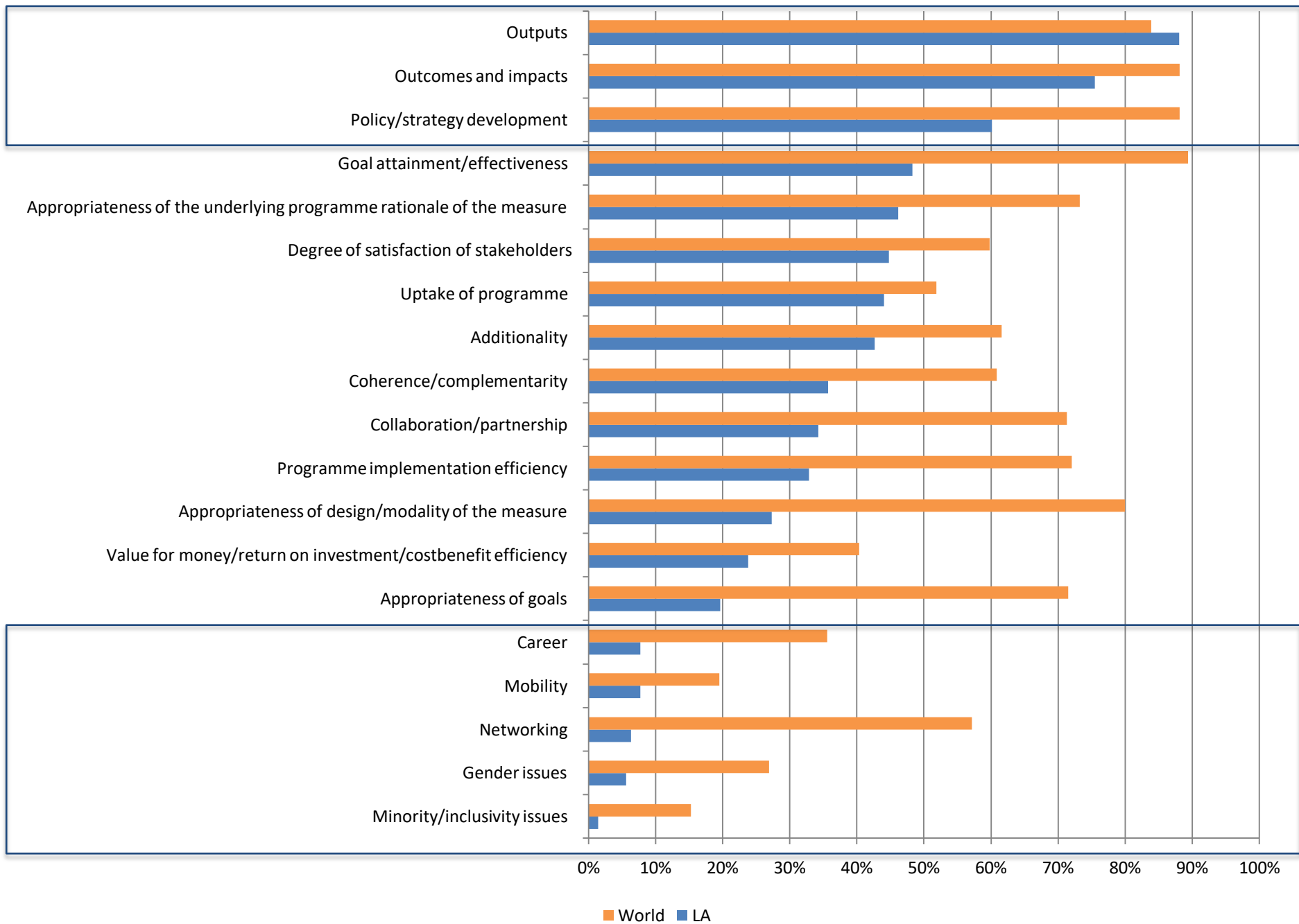
Timing	LA	World
Interim (periodic ex-post, after a specified phase during the implementation of the measure/programme)	125	401
Ex-post/final (after the lifetime of the measure)	12	80
Ex-ante (before the implementation of the measure/programme)	6	10
Accompanying (on a permanent or repetitive basis during the implementation of the measure/programme)	0	74
	143	565

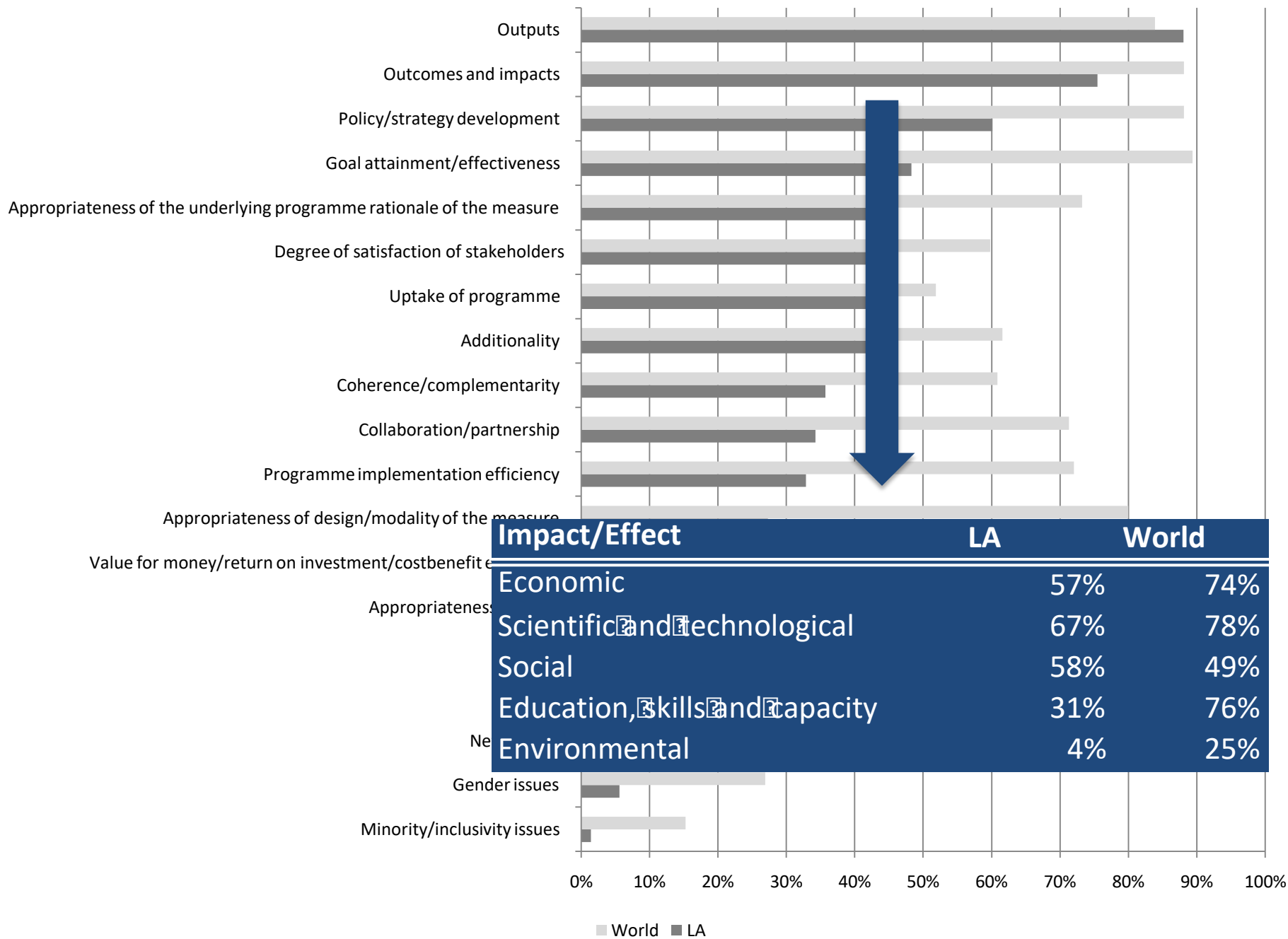
Purpose	LA	World
Both	73	392
Summative	54	102
Formative	16	66
Blanks	0	5
	143	565

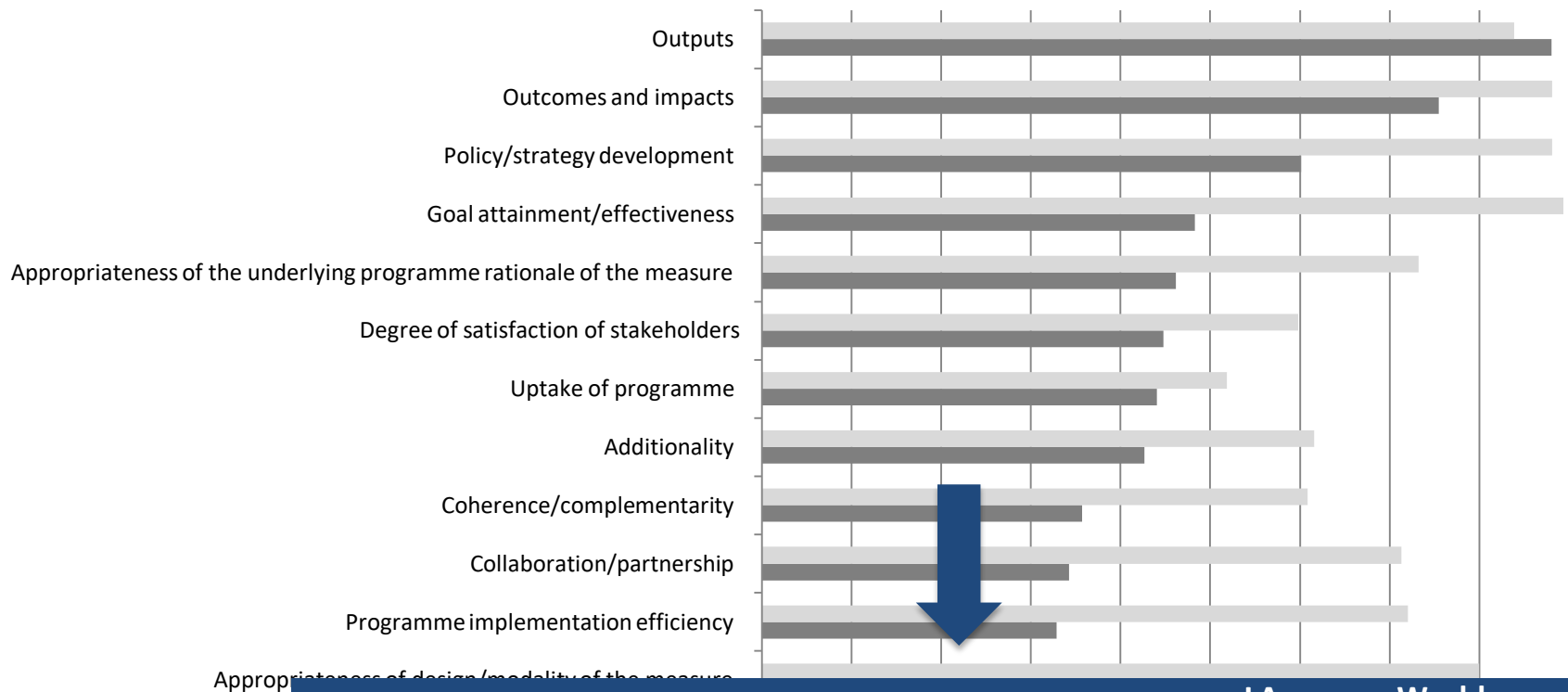
Reference to the programme logic or its intervention rationale	LA	World
Yes-fully (it clearly refers to the rationale for its development and identifies the way in which the intervention achieves the stated objectives (e.g. by using a logic chart model))	75	288
Yes-partially (it refers in a broad sense to the original rationale for establishing the programme/measure)	57	258
No	11	19
	143	565

MOSTLY INTERIM, FOLLOWED BY EX-POST (EX-ANTE EVALUATIONS ARE SCARCE OR NOT PUBLIC)
 MOSTLY SUMMATIVE AND FORMATIVE
 MOSTLY REFERRING TO INTERVATION RATIONALE (FULLY OR PARTIALLY)

**WHICH ASPECT OF THE PROGRAM DID
THE EVALUATION EXAMINE?**







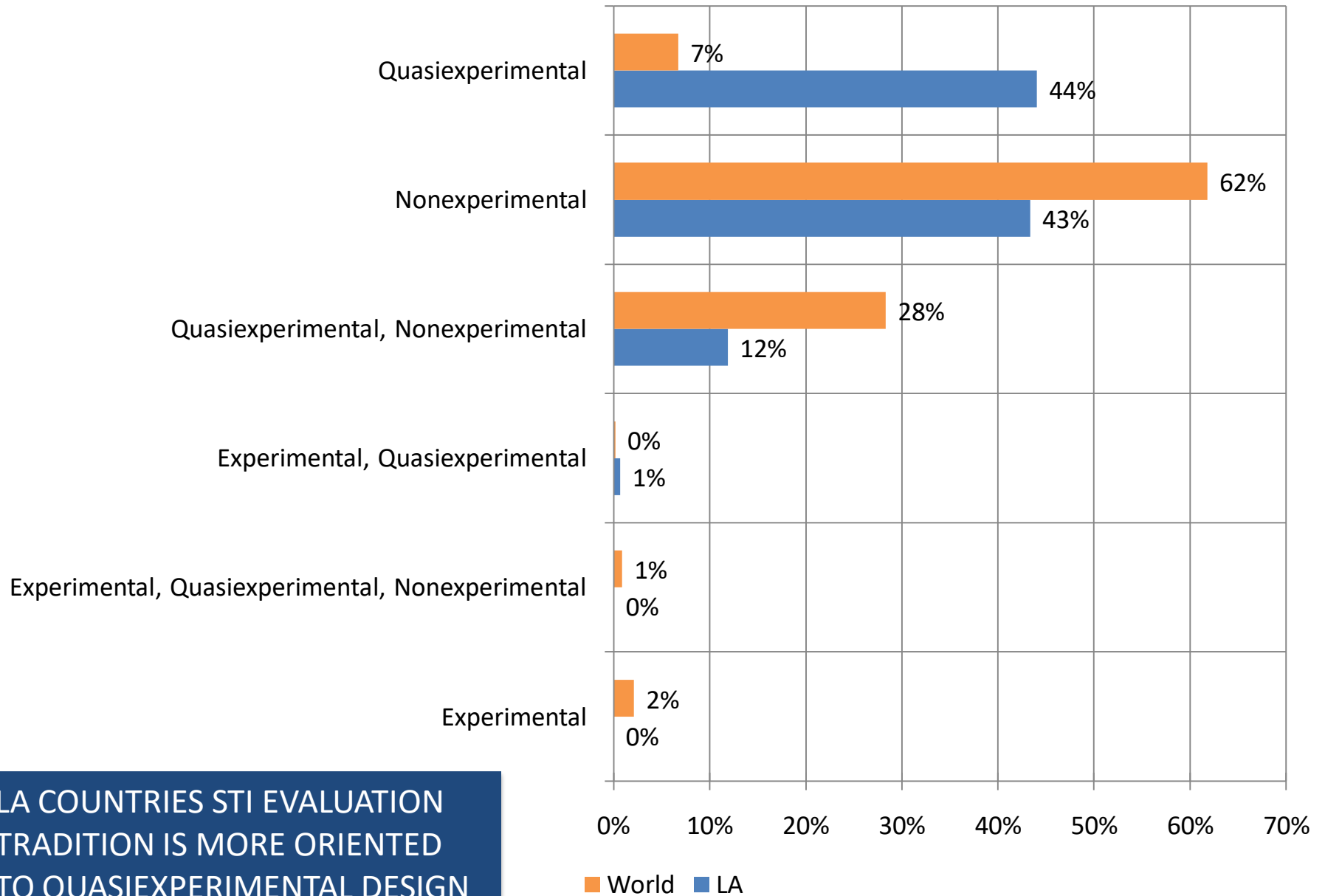
	LA	World
Value for money/ Input additionality (e.g. Does the evaluation examine if the measure stimulated more investment in RTDI than would have occurred in the absence of the measure?)	46%	67%
Output additionality (e.g. Does the evaluation examine if the measure stimulated more RTDI outputs than would have occurred in the absence of the measure?)	82%	81%
Behavioural additionality (Does the evaluation examine if the measure stimulated persistent changes in the behaviours of the participants that would have not occurred in the absence of the measure?)	15%	64%

100%

■ World ■ LA

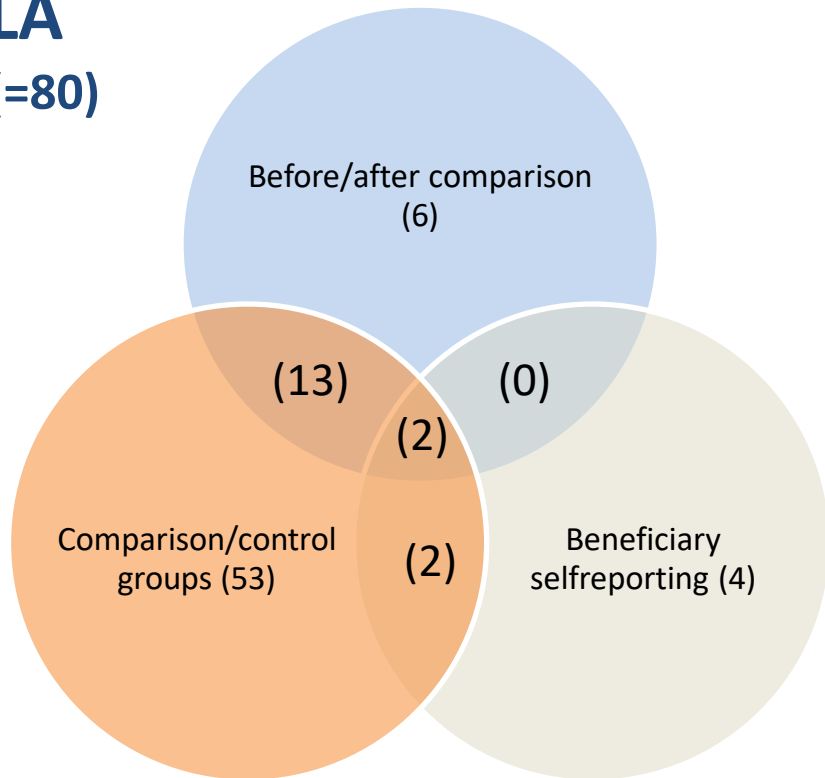
**WHAT ARE THE EVALUATION
DESIGNS?**

LA COUNTRIES STI EVALUATION TRADITION IS MORE ORIENTED TO QUASIEXPERIMENTAL DESIGN (influenced by IADB?)

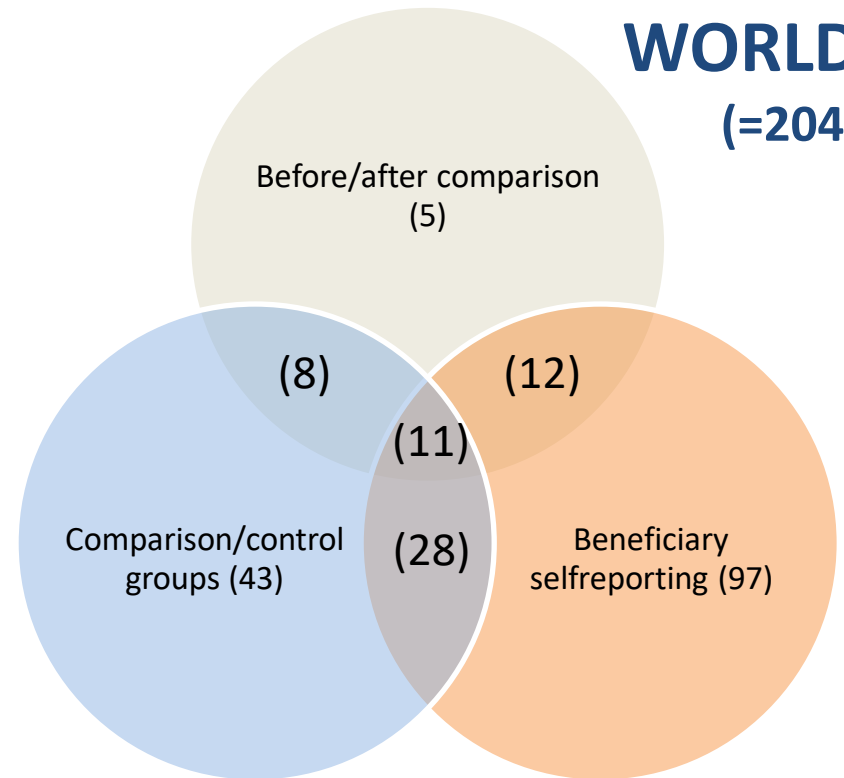


Design	LA	World
Quasiexperimental	63	38
Nonexperimental	63	349
Quasiexperimental, Nonexperimental	17	160
Experimental, Quasiexperimental	0	1
Experimental	0	12
Experimental, Quasiexperimental, Nonexperimental	0	5
	143	565

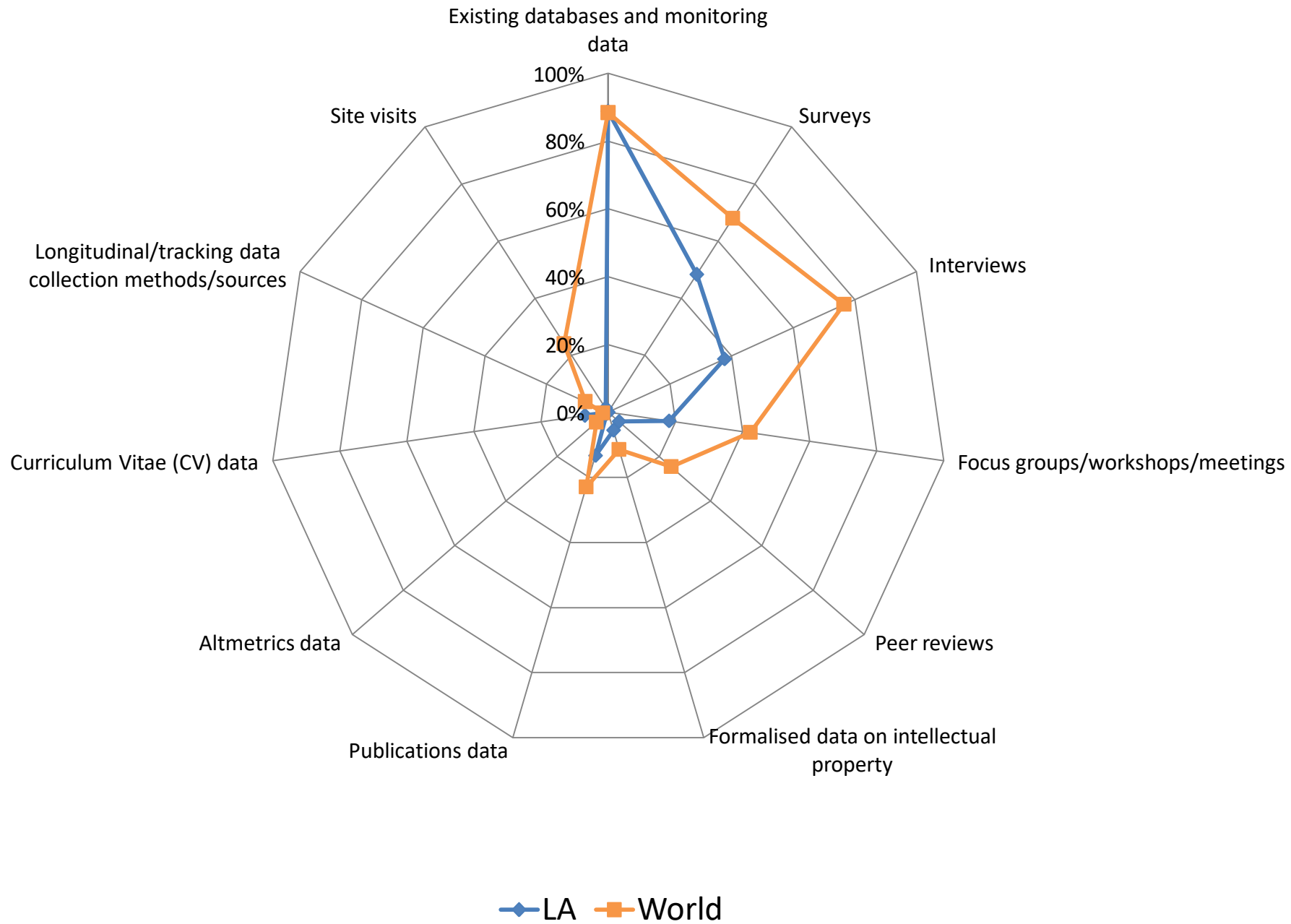
LA
(=80)

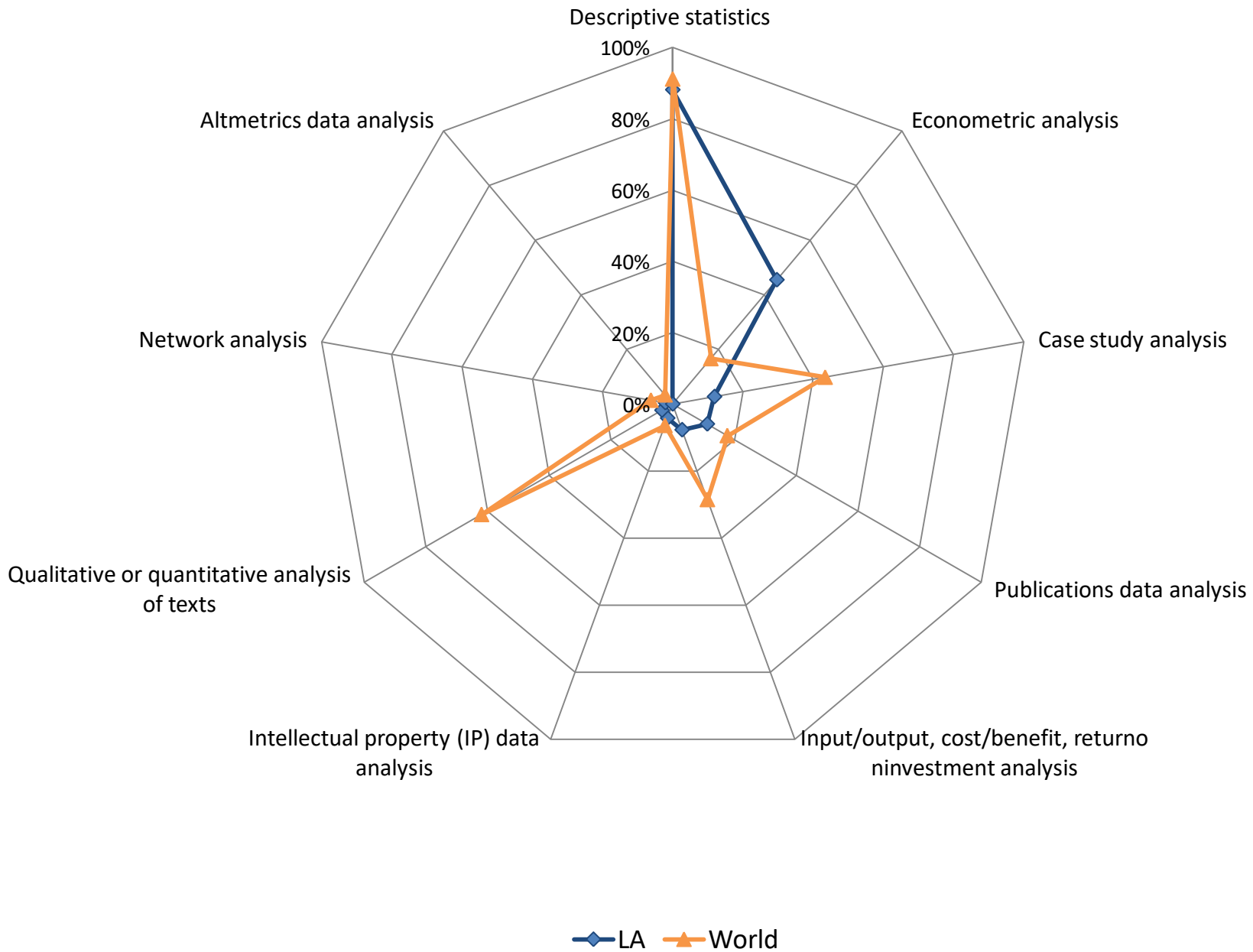


WORLD
(=204)



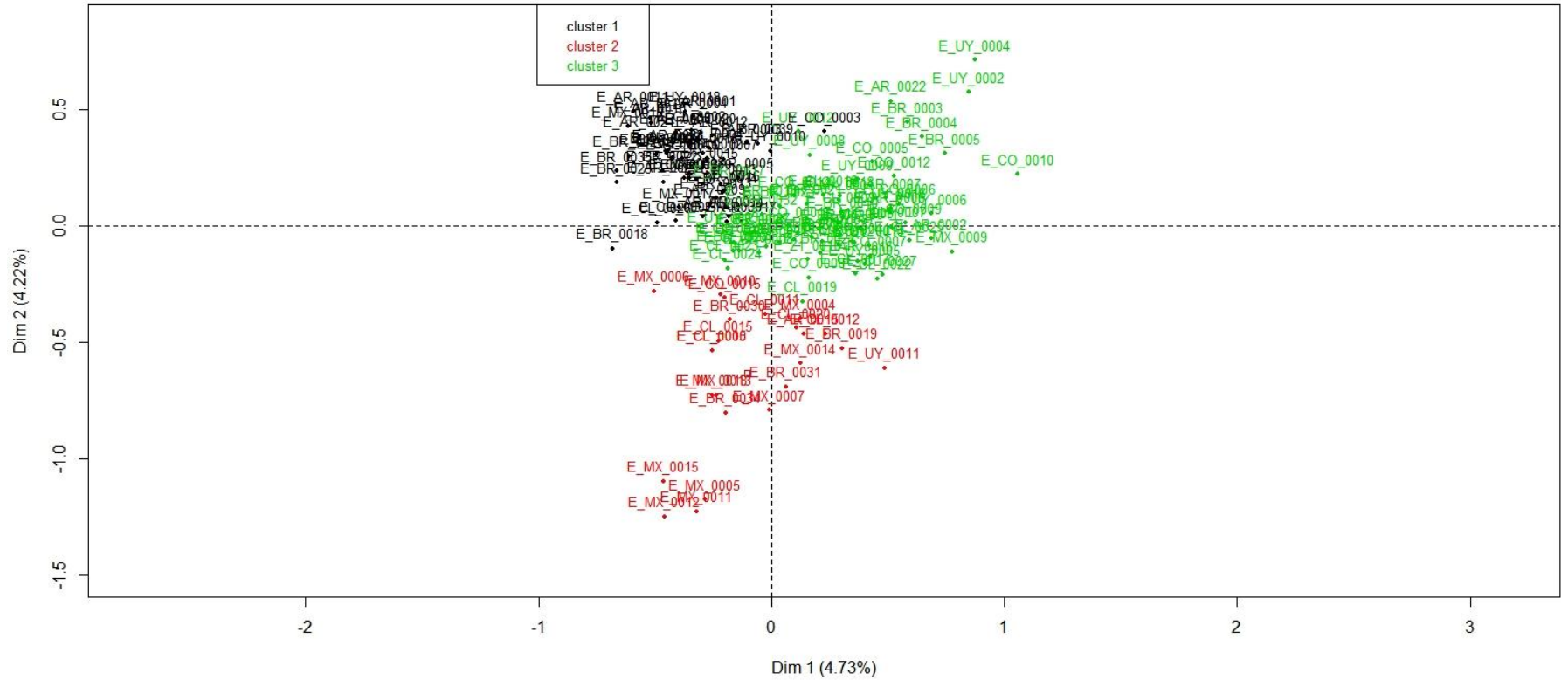
**WHAT ARE THE DATA COLLECTION
AND DATA ANALYSIS METHODS?**





CAN WE FIND SOME PATTERNS?

Factor map



	CLUSTER 1 (50)	CLUSTER 2 (24)	CLUSTER 3 (68)
Country	38% Argentina, 26% Brazil	46% México, 25% Chile	29% Brazil, 21% Chile, 19% Uruguay
Who_conducted	External independent	External independent	External independent and Internal
Purpose	Summative	Formative	Summative and Formative
Intervention_rationale	Fully	Partial	Fully
Appropriateness_rationale	Yes and No	Yes	No
Coherence_complementarity	No	Yes	No
Goal attainment/effectiveness	Yes and No	No	Yes
Ev_outputs	Yes	No	Yes
Ev_outcomes	No	No	Yes
Programme implementation efficiency	No	Yes	No
Policy strategy development	No	Yes	Yes
Uptake	No	No	Yes
Satisfaction stakeholders	No	No	Yes
Collaboration	No	No	Yes
Design	Quasiexperimental	Non experimental	Quasiexperimental and Non experimental
Surveys	No	No	Yes
Interviews_Workshop	No	Yes	Yes
Analysis	Econometrics and/or Descriptive	Descriptive only	Just Descriptive, Descriptive and others

Table 4. Qualitative summary of cluster profiling

Evaluation dimensions	Cluster 1: the support	Cluster 2: the verdict	Cluster 3: the holistic
Timing	Interim (68%)	Ex post (82%)	Ex post (75%)
Purpose	Formative (68%)	Summative (64%)	Both (70%)
Planning	Foreseen and planned (85%)	Less often foreseen and planned (46%)	Foreseen and planned (85%)
Conducted by	External (98%)	External (46%), but also internal (36%) or mixed (18%)	External (90%)
Topics	Programme efficiency (85%) (and thus management) focused, also consistency (83%/87%) coherence/complementarity (74%) and policy/strategy development (74%) important	Target few topics: mostly output (64%) and goal attainment (55%), also some input additionality (55%); not about internal (9%) or external (0%) consistency, project implementation efficiency (9%) or policy/strategy development (9%)	Target many topics: esp. goal attainment (100%), output (100%) and quality (80%); all types additionality (90%), but also consistency (external (80%)/internal (70%)), programme implementation and policy/strategy development (70% each)
Impact	Impact assessment important (89%), but only considers technological and economic impact (about 50% each)	Impact assessment less often used (64%) but still most important topic. Mainly economic impact (half of evaluations)	Clearly focused on impact (100%), all cover economic impact, 75% technological impact, 60% social impact
Methodological approaches and data sources	Qualitative methods and sources important; either interviews (94%) and focus groups (60%) or document (60%) and context (72%) analysis; participant surveys (77%) and existing surveys/databases (68%) used, but descriptive (79%)	Narrow approach, only few methods/sources used; mostly quantitative (econometric analysis (55%), control group (55%), counter-factual (64%)), based on existing data (46%) and participant surveys (46%)	Broader scope: many methods partially used, esp. interviews (90%), participant surveys (90%) and existing data (70%) are all important; but analysis is restricted to descriptive statistics (100%); In addition, case studies (40%), context analysis (55%), input/output (20%), group comparison (before/after; 10%), cost/benefit analysis (35%)
Measure type	Relative higher share of innovation diffusion (40%) and uptake measures (25%)	2/3 are about direct financial support (accountability)	Focus on science-industry cooperation (45%) network (35%) and spin-off (20%) programmes

Main findings (to LA) - I

- There is not an evaluation network (there should be?)
- Important role of research and innovation councils (institutionalization of evaluation efforts)
- Ex-ante evaluation is not a consolidated practice (or it is not public)
- Most evaluations looking at outputs and outcomes but few looking at education/skills, environmental issues, mobility, career, gender, minority issues (there is a gap in issues that are really important for developing countries)
- Measuring behavioral additionalty is something to be developed

Main findings (to LA) - II

- There is a need to further explore new evaluation designs, data collection and data analysis tools (this also means new insights from evaluations)
- There are some different traditions (not so much characterized by countries), reflecting different conditions and the background of evaluation practices – there is no need for convergence towards best practices, since “*one size doesn’t fit all*”)

WHAT IS AHEAD?

Agenda

- More evaluation collection, more characterization (\approx 200 documents)
- Include Policy modality to Correspondence and Cluster Analysis
- Quality assessment of evaluations (data already in the repository)
- Assessment of evaluation use (new data)
- Meta-analysis and Evaluation Synthesis (what are the overall effects of policies?)



THANK YOU

ADRIANA.BIN@FCA.UNICAMP.BR