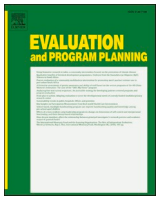
Evaluation and Program Planning 108 (2025) 102509

Contents lists available at ScienceDirect

Evaluation and Program Planning

journal homepage: www.elsevier.com/locate/evalprogplan



Anticipatory evaluation. How to incorporate an anticipatory technique into a theory-driven evaluation process. Results of application in a case study.☆

Mariangel Pacheco – Troisi [a,b,*,1], Mónica García-Melón [a,2], Fernando Jiménez-Sáez [a,3]

[a] Ingenio (CSIC – Universitat Politècnica de València), València 46011, Spain
[b] Universidad Tecnológica (UTEC), Montevideo 11000, Uruguay

ARTICLE INFO

Keywords:
Evaluation
Strategic learning
Impact
Anticipation
Science
Technology
And innovation

ABSTRACT

In recent years, there has been increased focus on strategic learning from impact analysis, including in the field of science, technology, and innovation. In this paper, we propose combining techniques from two fields of study. Firstly, we adopt the approach of addressing impact through productive interactions between science and society, and secondly, we incorporate an anticipatory dimension by integrating game analysis involving key actors. Through a theory-driven evaluation design, we consider expected impacts as promises of the future. Within an anticipatory perspective, the future can be shaped by the interactions among different actors in the present. In this article, we apply this approach step-by-step to a research institute program in Uruguay. We demonstrate how the achieved results offer strategic insights to the program manager for anticipating and attaining the desired impacts. Additionally, we provide summative inputs for accountability using a flexible technique applicable at any stage of the program life cycle. The article concludes with a discussion of the advantages and disadvantages compared to other techniques, along with lessons learned that may benefit other evaluators seeking to replicate this approach. Furthermore, we explore potential extensions and opportunities for further improvement in this research.

1. Introduction

Impact evaluation in the field of Science, Technology and Innovation (STI) is an area where new forms of evaluation design and practice are required (Matt, Gaunand, Joly, & Colinet, 2017). As proposed by Wolf, Lindenthal, Szerencsits, Holbrook, & Heß (2013) methodologies with greater emphasis on the evaluation of preconditions, intermediate steps to achieve the expected impacts.

Traditional evaluation approaches have been limited in their ability to respond to growing expectations about the capacity of STI to generate impacts on society. These approaches use reductionist indicators (Ràfols, 2018), and have classic problems of temporality (Buxton, 2011) as well as a lack of understanding of the mechanisms and processes for achieving impact (Molas-Gallart & Tang, 2011). At the same time, increasing emphasis is being placed on the growing variety of users (Castro-Martínez, Olmos-Peñuela, & Fernandez-de-Lucio, 2016) and uses of science (Cozzens & Snoek, 2010). Furthermore, current science policies that aim to enhance the social value of research are often concerned with the relationships between scientific and social actors (Smit y Hessels, 2021). At the end of the day, this also translates into increased pressure on researchers, who face the challenge of fulfilling their scientific and social mission at the same time (D'Este, Ramos-Vielba, Woolley, & Amara, 2018). Although there are authors with a constructivist stance who claim that the social and scientific value of research are strongly linked (Smit & Hessels, 2021), it is by no means well established that scientific excellence is an adequate predictor of social value (Buxton, 2011).

These criticisms have led to the emergence of a significant number of projects that have in common the idea of extending the impact capacity of science rather than reducing it to measurability. This family of projects is known as RIA (Research Impact Assessment). One of the milestones in this organisation of research towards the grand challenges of

☆ This work has been partially funded by Fundación Carolina.
* Corresponding author at: Ingenio (CSIC – Universitat Politècnica de València), València 46011, Spain.
E-mail addresses: mapactr@doctor.upv.es, mariangel.pacheco@utec.edu.uy (M. Pacheco – Troisi).
[1] ORCID: 0000-0003-1932-1851
[2] ORCID: 0000-0002-4587-8920
[3] ORCID: 0000-0001-5685-3959

https://doi.org/10.1016/j.evalprogplan.2024.102509
Received 6 June 2023; Received in revised form 8 December 2023; Accepted 27 September 2024
Available online 16 October 2024

STI assessment has been the Lisbon Agenda (2001) (P. B. Joly et al., 2015).

We take up the results of the *Social Impact Assessment Methods for research and funding instruments through the study of Productive Interactions between science and society* (SIAMPI) project, which proposes that science can only generate impact if there are productive interactions (PIs) between science and society (Spaapen, Shinn, & Marcovich, 2012) and these are a transparent approximation of the process from research to impact (Manrique, Wróblewska, & Good, 2019). PIs can be considered predictors of successful impact (P. B. Joly et al., 2015) and could have a trajectory, a time horizon associated with them (Damşa, 2014). That is, they could have an evolution from moments of momentary interaction (Krange, 2007). In addition, the characteristics of these PIs are also important, i.e. their typology (Spaapen et al., 2012), variety of actors involved, breadth and depth, and presence of interactive learning processes (D'Este et al., 2018).

On this fundamental background, we add an anticipatory dimension, which implies the explicit recognition that it is the decisions made in the present that create the future (Martin, 2010). There is not one but many possible futures and at any given moment the future is multiple and it is from the confrontation of the different actors that we will end up with one future rather than another (Godet, 1991).

We consider the MACTOR method (*Méthode ACTeurs, Objectis, Rapports de force*), a specific method for the analysis of the actors' game, which makes it possible to estimate the correlations of forces that exist between them and to study their convergences and divergences in relation to certain strategic objectives (Durance, 2011).

Stufflebeam (2001) classified more than 22 streams of evaluation, demonstrating that doing evaluations involves making choices and options from a wide range of possibilities. On this basis and in a context where there is interest in formulating and applying new ways of addressing the social impact of STI, in this paper we propose to combine resources from two different disciplines. For that, we argue how to incorporate an anticipation technique into a theory-driven evaluation process and demonstrate this assumption for a particular case study. This new approach is called Anticipatory Evaluation (AE) and consists of 3 approximative phases including 10 steps.

The main aim of AE is to provide information that contributes to the strategic vision of STI project managers and contribute to their decision-making processes. It also aims at identifying mechanisms to dynamize the trajectory of productive interactions with various users to anticipate and achieve the expected impact of their projects. It is a formative technique, which enables strategic learning from impact analysis and combines quantitative and qualitative resources. It focuses on individual analysis, which allows working with different levels of aggregation.

In this paper we will show the step-by-step application of this technique in a research institute in Uruguay. For this purpose, the rest of the paper is structured as follows: in section II we present the methodology, in section III we present the step-by-step application and its results, in section IV we present a discussion on the advantages and disadvantages in comparison to other evaluation approaches. In doing so, we focus on the classification (Stufflebeam, 2001) and on an own systematisation of impact assessment methods specifically in the field of STI. Finally, in section V we present a synthesis of lessons learned that may be useful for other evaluators interested in replicating this work, as well as opportunities for improvement that could be exploited.

2. Methodology

This article is a contribution to strategic learning processes based on impact analysis. We build upon the productive interactions approach, resulting from the SIAMPI project (Spaapen & van Drooge, 2011) by proposing a methodology that prioritises learning and transparency in processes to achieve impact and adds an anticipation dimension using the specific method of the actor game from the strategic foresight toolbox. This proposal consists of 10 steps grouped into 3 stages, as illustrated in the flowchart in Fig. 1.

In the first stage of an AE process, the evaluation team (researcher and evaluator) is formed and the evaluation unit is defined using the logic model or programme theory (Alkin, 2011). It should be clarified that the term researcher is used to refer to the role of the STI programme leader/manager and it is assumed that the name of this role may vary in other cases. In the second stage, work is done to identify the broad set of actors with whom it would be necessary to establish PIs to achieve the expected impacts according to the programme theory. These actors are classified according to the environment to which they belong (private sector, public policies, basic research, applied research, etc.). Existing PIs are then identified and classified according to their typology, depth, and presence of bidirectionality. Those PIs that are pending are also identified with those actors that have been identified as key to achieving the committed impacts but that the researcher has not yet engaged with. In short, in this second stage a picture is built for a given moment of the existing and pending productive interactions and their characterisation. In the third stage, the MACTOR method of the actors' game is applied to identify the convergences and divergences that exist between them in relation to the strategic objectives of the evaluated project. This third stage concerns the future, and its purpose is to provide the project leader with information for strategic decision-making and to anticipate to dynamize the trajectory of existing PIs or to initiate pending ones. In other words, to confront the different actors to move towards the desired future, the one identified in the programme theory. To do this, the leader/manager needs very specific information that allows him/her to approach those actors with whom he/she has not yet been able to establish links or whose links are stagnating. Semi-structured interviews are used to learn about a specific problem concerning a key actor and to identify how the intervention can contribute to a solution. In other words, an effective connection between problems and solutions.

Our proposal, presented in Fig. 1, is based on the participation of people with three clearly differentiated roles: the researcher, the evaluator, and the key actors. For each of them and for each step of the methodology in particular, different tasks are assigned to them.

3. Implementation and Results. Case study in the rice research programme in Uruguay

3.1. Stage I - identification of the programme leader, the evaluator and the evaluation unit

3.1.1. Steps 1 and 2. The programme leader, the evaluator, and the evaluation unit

The first step is the setting up of a team composed of the researcher (programme leader) responsible for the programme and the evaluator who will lead the evaluation.

Evaluation methods are not passive (De Rijcke, Wouters, Rushforth, Franssen, & Hammarfelt, 2016). Nor, we believe, is the role of the evaluator. We therefore assign a vitally important role to the evaluator, with a committed approach that proposes evaluation as an analytical and strategic project from a critical change-oriented perspective (M. Q. Patton, 2002) or transformative paradigm (Ward Hood & Cassaro, 2002).

In our case study the unit of evaluation is the Rice Research Programme (RRP) of the Uruguayan research institute *Fundación Latitud*. Fig. 2 describes the objectives and Fig. 3 presents the illustration of its logic model (Alkin, 2011). In terms of previous organisers (Guba y Lincoln, 1989), AE is a methodology that works on a theory-driven evaluation design and implementation (Chen, 1990).

At this point, we consider it necessary to specify the evaluation questions (Ligero Lasa, 2015) that guided this methodological design.

Has the broad set of actors with whom the PIs need to be established to achieve the intended impacts of the programme been identified?

Is the coverage of existing PIs adequate for the life cycle phase of the programme?

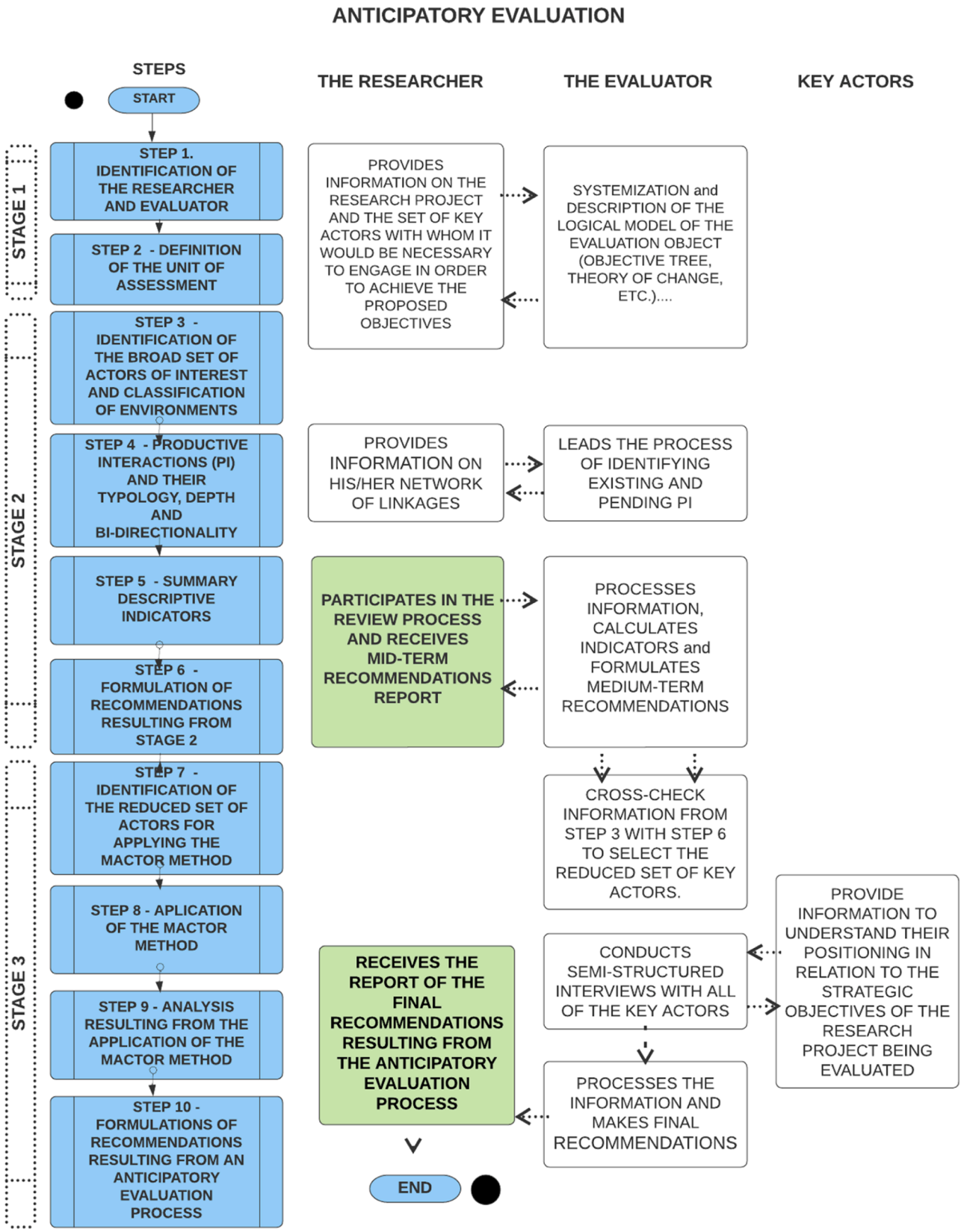


Fig. 1. - Flowchart - Anticipatory Assessment. Own elaboration.

Is the typology, depth and bidirectionality of existing PIs appropriate for the programme life cycle phase?

Is there an adequate plan (strategic planning) to initiate pending PIs?

Is there an adequate plan (strategic planning) to activate the trajectory of existing PIs?

Object of Evaluation	Rice Research Programme (RRP) of the Uruguayan applied research institute Latitud foundation.
General objective	Improving the competitiveness and sustainability of the Uruguayan rice sector.
Specific objectives	1) Increase industrial profitability by optimising industrial processes: optimising industrial drying yields, optimising milling yields and minimising quality loss during storage. 2) Contribute to increasing the competitiveness and sustainability of the rice value chain by incorporating new rice varieties into the production system, thus accelerating the exploitation of the genetic gain of the new cultivars and increasing the efficiency of each link in the chain. 3) Contribute to guaranteeing the safety of Uruguayan rice production, anticipating the eventual incorporation of strict requirements in terms of safety standards.

Fig. 2. - Descriptive presentation of the RRP.

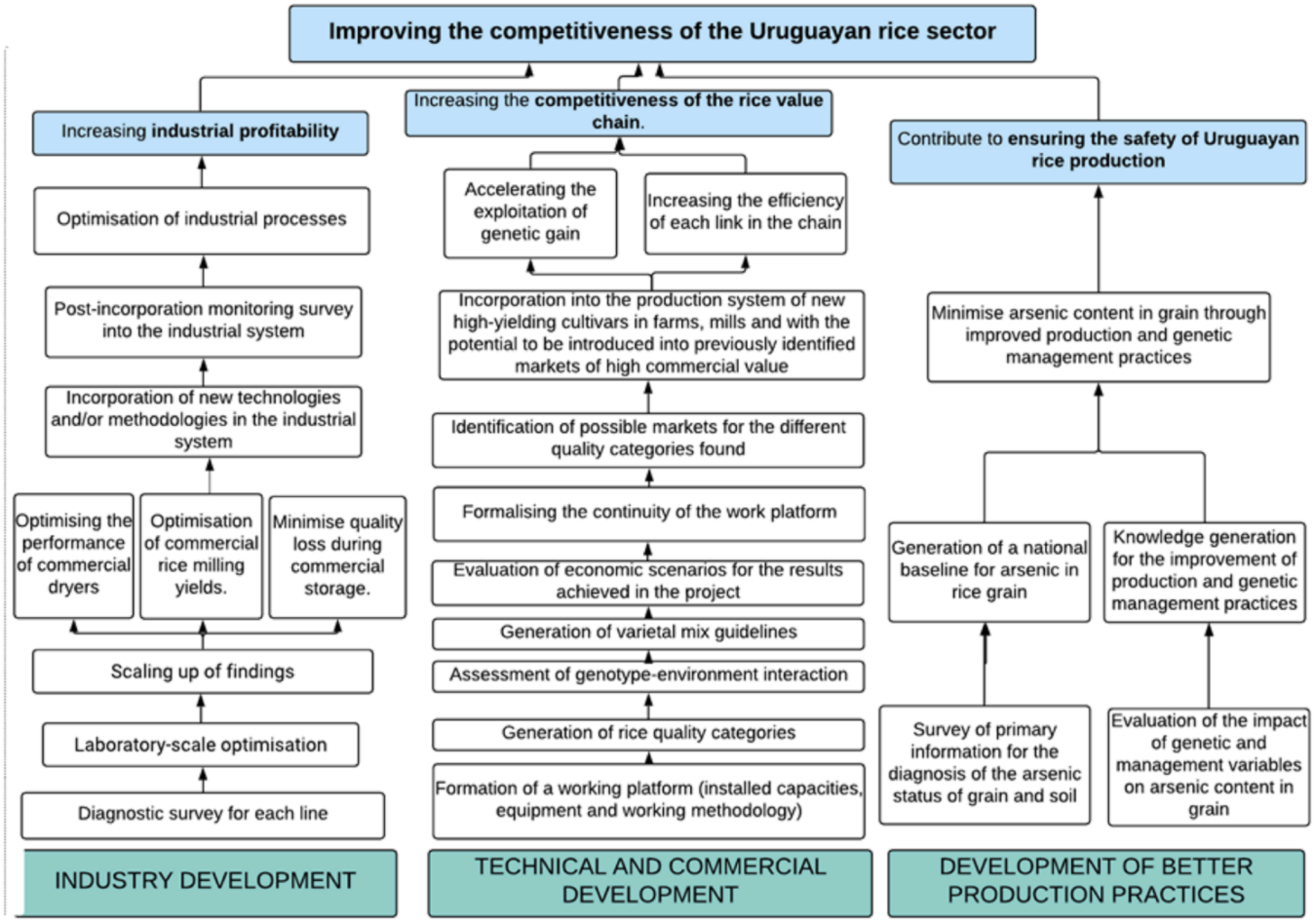


Fig. 3. - RRP logic model.

3.2. Stage 2 - existing and pending productive interactions (PIs)

3.2.1. Step 3. Identification of the broad set of actors of interest and classification of environments

Research relationships with social actors differ, even within the same research field (Hessels, van Lente, Grin, & Smits, 2011). As a first step in identifying existing and pending PIs, the set of actors with whom the programme leader has already established links or foresees the need to do so to achieve the expected final impacts of the programme, is identified. This process is guided by the evaluator, answering the question: to fulfil this component of the logic model, which actors should be involved? This is an information triangulation exercise. On the one hand, the evaluator analyses the background of the ecosystem of actors relevant to the programme. The programme leader (the researcher), based on his or her specific experience and expertise, also proposes key actors. Finally, the evaluator cross-checks with publicly available information regarding the actors, organizations and institutions that were identified in the process. As an example, mission, purpose and scope statements on official websites. In short, we can say that behind the logic model, there is also a tree of actors with whom it is necessary to engage

to achieve the expected final impacts. For this case study, 38 actors were identified, as illustrated in Fig. 4.

We cannot guarantee the success of the programme by having PIs with all identified actors, just as we cannot guarantee that all key actors have been identified. However, we know that the more productive interactions between researchers and society, the higher the expectations of social impact (Maassen van den Brink et al., 2010) and therefore the PIs can be considered predictors of project impact (P. B. Joly et al., 2015)

Each of the identified actors are classified according to the environment to which they belong: i) private sector, ii) applied research, iii) basic research, iv) other organisations and v) public policy (Spaapen et al., 2012). They are also classified by geographical environment: (i) national, (ii) regional and (iii) international. In supplementary material 2 and 3 we present the classification according to environment and geography of the broad set of actors for this case study.

3.2.2. Step 4. PIs and their typology, depth and bidirectionality

A productive interaction (PI) is a link between a stakeholder and the project. This relationship has to involve some effort by that stakeholder to engage with the project (Spaapen et al., 2012). In this step, work is done on recording existing PIs and identifying potential or pending PIs.

Regarding existing PIs, authors such as (Sivertsen & Meijer, 2020) stress the importance of observing daily, active, productive and receptive relationships. In our proposal we work on the conformation of two subgroups of PIs: existing and pending. We first analysed the existing subgroup and classified them according to the following categories and subcategories presented in Fig. 5 (Spaapen et al., 2012), (D'Este et al., 2018).

All this information is processed with the logic of a database. In our case, we work with the statistical programme SPSS (*Statistical Package for Social Sciences*). Supplementary material 4 is the database (.sav).

3.2.3. Step 5. Summary of descriptive indicators

In this step, the evaluator calculates a set of indicators that synthesise the previous steps and precede the formulation of the intermediate recommendations resulting from the second phase. These indicators are: i) total number of identified actors and their distribution by environment and geography, ii) coverage indicator, typology of existing PIs and their distribution by environment and geography, iii) intersection of typology, depth and bidirectionality of existing PIs, iv) existing versus pending PIs by environment.

1) Total number of actors identified and their distribution by environment and geography

42 % of the 38 actors belong to the private sector, 18 % to the public policy environment, 16 % to the group of other organisations, 13 % to the field of applied research and 11 % to basic research.

76 % of the actors belong to the national level, while the rest are equally distributed between the regional and international spheres, with a share of 13 % and 11 % respectively.

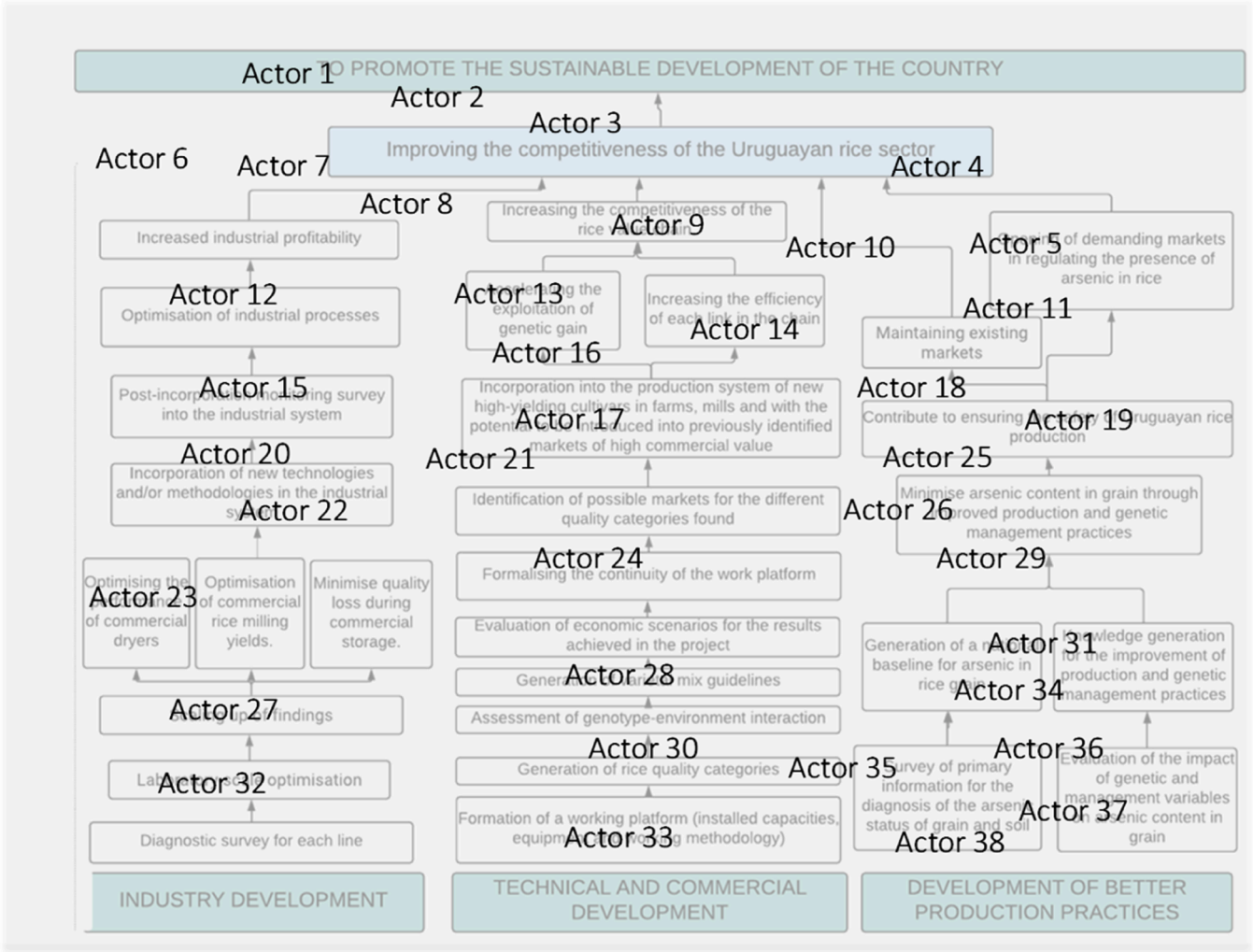


Fig. 4. - Set of actors of interest to the RRP. This image is only intended to illustrate the process of identifying the broad set of actors.

Category	Subcategories
Typology	Type I: direct or personal
	Type II: indirect or through some medium or material
	Type III: financial
Depth [1]	Low
	Medium
	High
Presence of bidirectionality [2]	1: when there is a specific verifier of this bidirectionality between the researcher and the actor.
	0: when there is no verifier of bidirectionality.

Fig. 5. - Categories and sub-categories for the classification of existing PIs.

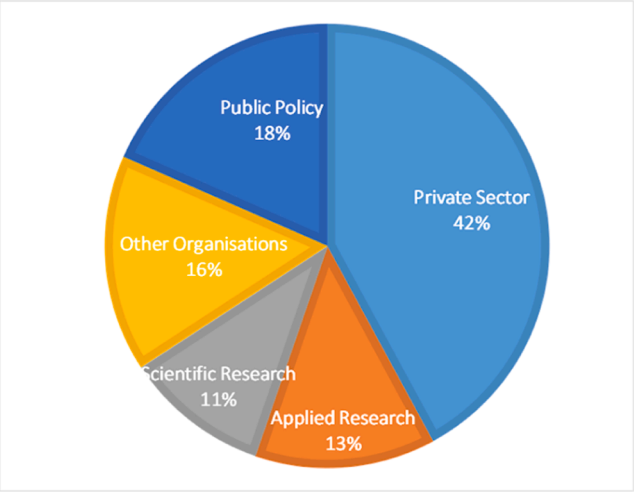


Fig. 6. - Distribution by category of environment. Total number of actors: 38. Case study: RRP.

GEOGRAPHICAL DISTRIBUTION		
National	Regional	International
76%	13%	11%

Fig. 7. - Distribution by category of geographical affiliation. Total number of actors: 38. Case study: RRP.

2) Coverage indicator, typology of existing PIs and distribution by environment and geography.

The programme team has established PIs with 22 of the 38 key actors, which represents a coverage indicator of 58 %.

59 % of these interactions are type III (financial), 27 % type I (personal) and 14 % type II (through some artefact or material).[4]

More than half of these actors (54 %) belong to the private sector. The remaining 10 actors are similarly distributed between applied research, basic research, and other organisations.

These 22 actors are mostly from the national level (73 %), followed

[4] This classification is not exclusive, and a PI can simultaneously belong to more than one category. In this case, through a process of dialogue with the researcher, the most influential typology is identified.

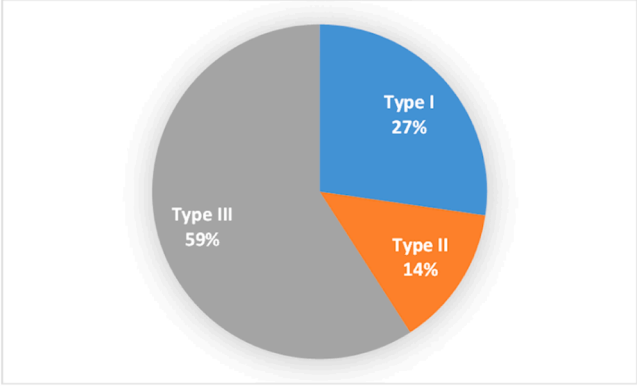


Fig. 8. - Distribution of existing PIs according to typology. Total number of actors: 22. Case study: RRP.

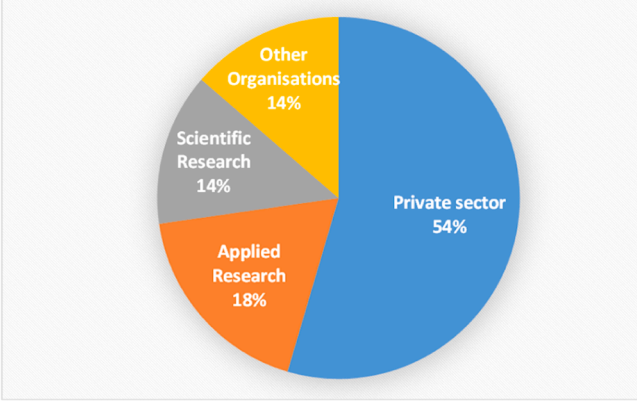


Fig. 9. - Distribution of existing PIs by environment. Total number of actors: 22. Case study: RRP.

CLASSIFICATION OF ENVIRONMENTS		
National	Regional	International
73%	18%	9%

Fig. 10. - Geographical distribution. Total number of actors: 22. Case study: RRP.

by the regional and international levels (18 % and 9 % respectively).

3) Intersection of typology, depth and bidirectionality of existing PIs

Of the 22 PIs, 10 (47.6 %) have been identified with bidirectionality. Of these actors, 9 belong to the private sector and 1 to the category of "other organisations".

All 10 cases of bidirectionality are found in PIs of the third type, those interactions involving financial aspects. Within these, they are almost equally distributed (3, 4, 3) for the low, medium, and high depth levels, respectively. Of the remaining 12 PIs, those without bidirectionality, we find that 3 are type III, of which 2 are of medium depth and 1 of high depth. 3 are type II with shallow and medium depth and 6 are type I with low depth.

4) Existing and pending PIs by environment

Fig. 11 shows that the most balanced distribution between existing and pending PIs is in the "other organisations" environment, where PIs have been established with half of the identified actors. For basic research and the private sector, we find that PIs have already been initiated with 75 % of the key actors. In the case of the applied research environment, existing PIs represent 80 % of the total number of PIs to be achieved.

Finally, it is worth noting what is happening in the public policy environment, where at the time of implementation none of the interactions needed to achieve the impacts committed to by the RRP have been initiated.

3.2.4. Step 6. Formulation of recommendations resulting from PHASE 2

In this step, the evaluator formulates recommendations (Carol H. Weiss, 1993) resulting from this strategic learning process based on the impact analysis (Melloni, Pesce, & Vasilescu, 2016). With the purpose of contributing to the decision-making process of the programme leader, with the evolution of the trajectory of existing PIs and the initiation of pending ones.

		Pending	Existing
Environment	Private Sector	25%	75%
	Applied Research	20%	80%
	Scientific Research	25%	75%
	Other Organisations	50%	50%
	Public Policy	100%	0%

Fig. 12. - Comparison - in percentage - of existing and pending PIs. Total number of actors: 38. Case study: RRP.

3.3. Stage 3: analysis of the actors' game and final recommendations of an ae process

Planning can be defined as the conception of a desired future and the means necessary to achieve it (Ackoff, 1973). Strategy can be defined as the set of rules of conduct of an actor which enable him/her to achieve his/her objectives and projects. Strategic foresight is a concept from the 1990s in which anticipation is put at the service of the project's strategic action. In turn, it can be distinguished between an exploratory phase - which tries to identify the challenges of the future - and a normative phase - which tries to define the strategic options to meet these challenges (Godet & Durance, 2007).

The incorporation of strategic foresight in this proposal corresponds to a normative phase in which the objectives, purposes and strategic impacts defined in the programme theory are to be achieved. The logical model of the object of evaluation corresponds to the future-producing intentionality of which Godet and Durance speak. These same authors propose strategic foresight by scenarios and define a scenario as a set consisting of a description of a future situation and a path of events that allow us to move from an original situation to a future one. The scenario method is a logical modular approach that does not necessarily have to be followed from beginning to end. It all depends on the degree of knowledge of the system under study and the objectives pursued. Depending on the needs, it is possible to limit the study to one or another module (Godet & Durance, 2007).

We propose to apply the module of actor game analysis (MACTOR method), which aims to assess the power relations between the actors

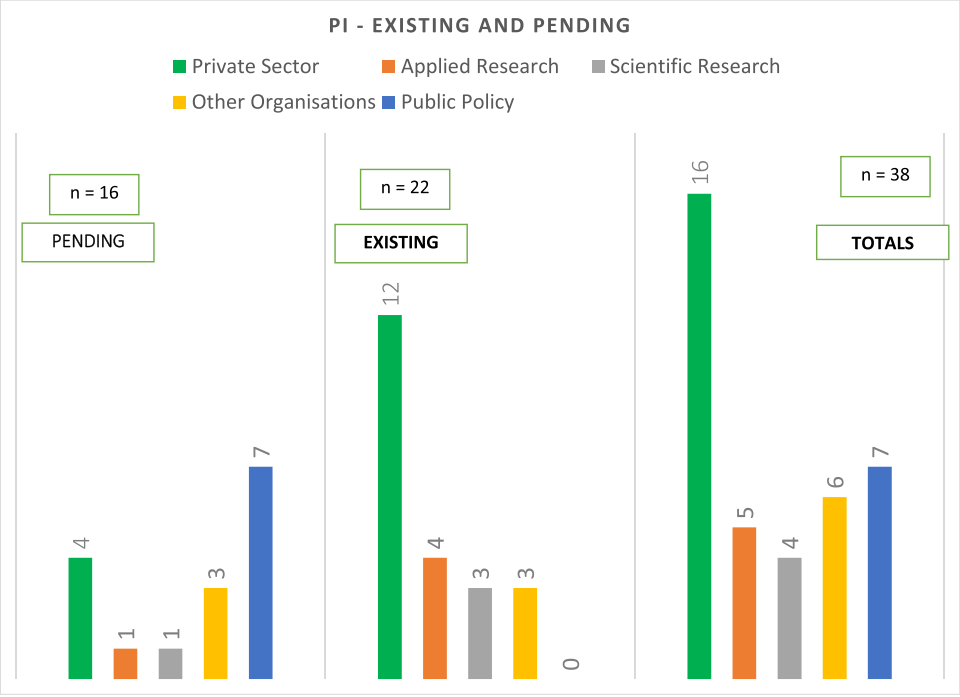


Fig. 11. - Comparison - in number - of existing and pending PIs. Total number of actors: 38. Case study: RRP.

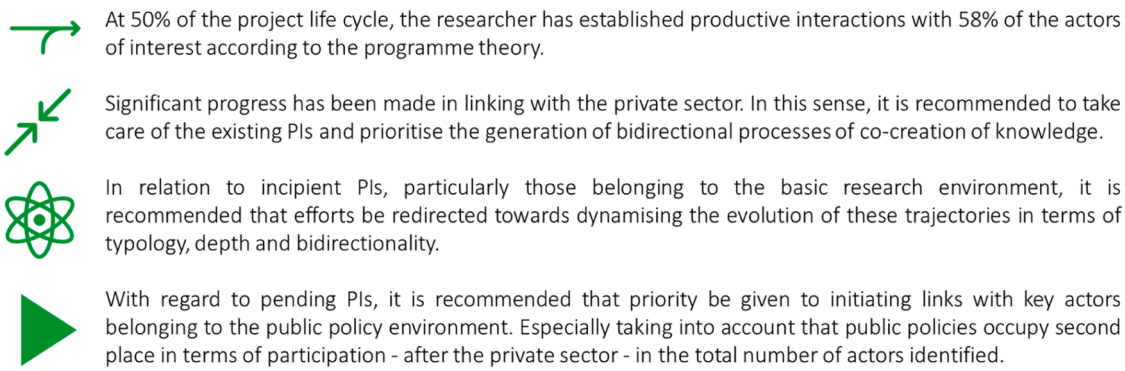


Fig. 13. - Synthesis of recommendations resulting from stage 2.

and to study their convergences and divergences with respect to a certain number of associated strategic objectives. Based on this analysis, the objective of using the MACTOR method is to provide an actor with information for the implementation of its strategy to reinforce convergences and attenuate divergences.

3.3.1. Step 7 - Identification of the reduced set of actors to apply the MACTOR method

In this step, the key actors that condition the implementation of the recommendations of step 6 are identified. Priority is given to those recommendations that refer to pending and incipient PIs. Fig. 14 shows, for the case study, the cross-referencing of prioritised recommendations, the identified elements that condition the initiation of the linkage or the dynamisation of the PI trajectory and the key actors that condition the evolution of the system of interest. These actors make up the reduced set and are the actors with whom the MACTOR method will be applied in the next step.

3.3.2. Step 8 - applying the MACTOR method

The application of the MACTOR method follows the steps detailed in (Durance, 2011). These are: i) construction of actor files and application of semi-structured interviews, ii) identification of the correlation of forces between actors in terms of power relations, iii) identification of strategic challenges and associated objectives, iv) positioning of the actors in relation to the objectives and identification of convergences and divergences (simple positions), v) prioritisation of objectives for each actor (evaluated positions), vi) integration of the correlation of forces in the analysis of convergences and divergences between actors.

These steps are automated in the free software MACTOR developed by LIPSOR (*Laboratoire d'Investigation en Prospective, Stratégie et Organisation*).[5]

The main results of this step are illustrated in Figs. 15 and 16. Fig. 15 illustrates these coincidences in the form of convergence nodes where most of the actors are located. The differences between the actors are to be found in the hierarchies that the interviewees assign to these objectives, and this has to do with the (indirect) compatibility they see with the strategic objectives of the institution in which they work.

The actors who agree most with their hierarchy of objectives are actors 1, 3 and 9. This is explained by the hierarchy they assign to the strategic objective called INOCTUITY. Actors 9 and 3 agree in assigning the first order of importance to this objective and specifically to the problem of the presence of arsenic in rice. In the same sense, the researcher from the basic sciences laboratory (actor 9) clearly identifies the convergence of the strategic objectives, particularly on the minimisation of the contaminant content in the grain, through the practice of productive and genetic management, and mentions the specific technique they are developing to evaluate the phenotypic characteristics of rice that behave as genetic indicators. This type of information could contribute to the process of incorporating new varieties of the grain. This is also linked to the strategic objective VALUE CHAIN, which refers to increasing the profitability of the entire rice value chain. Here, the official of actor 3 and the researcher from actor 9 point out the clear convergence in the objective that refers to the incorporation of new high-yielding cultivars into the production system in farms and mills, with the potential to be introduced in previously identified markets. It is precisely in the objective VALUE CHAIN where there is a greater convergence with actor 1, as the competitiveness of the whole rice value chain is a task of its public policy office. In relation to INOCTUITY, the interviewee identifies a clear convergence. However, he mentions that it is another area that works on this issue. So, in this second node of convergence, we can identify the actors: 2, 8, 4 and 7. The convergence between the first three is mainly due to the energy issue and how these impacts on the first components of the rice value chain and the subsequent effects on the efficiency of each link in the chain.

When the power relations between the actors are integrated, we can see in Fig. 16 that the map of convergences is reinforced. In particular, the convergence node between actors 3, 9 and 1 is reinforced, as well as the convergences of actor 3 with actor 4, actor 8 and actor 7. However, we can also see that actor 4 is still perceived as a relatively isolated actor, which is also evident in the interview with the technical staff of this institution.

From a graphical point of view, actor 2 could also be considered an actor with relatively weak convergences - except for the convergence with actor 3 - but the qualitative research revealed a clearly strategic and innovative positioning on the part of the technical staff, which is considered when formulating recommendations.

3.3.3. Step 9 - synthesis of the analysis of the results obtained by the MACTOR method

This step synthesises the main results of the analysis conducted in the previous step. This step is a bridge that connects the recommendations of step 6 - resulting from the analysis of existing and pending PIs - with the final recommendations (step 10) resulting from an anticipatory evaluation process.

3.3.4. Step 10 - Final recommendations resulting from an anticipatory evaluation process

Finally, in this last step, recommendations are formulated - complementing the recommendations of step 6 - with special emphasis on the analysis of the actors' game in the context of an AE process in which it is intended to contribute to the strategic perspective and decision-making process of the researcher in relation to the development and strengthening of her PI network.

[5] Open access at: http://www.laprospective.fr/methodes-de-prospective.htmlDetails of the step-by-step application of the MACTOR method for the case study are presented in supplementary material 5.

Recommendation from step 6	Elements identified by the researcher as conditioning factors in the evolution of the PI's trajectory	Actors with whom the MACTOR method is applied
Regarding the pending PIs, it is recommended that priority be given to initiating engagement with key actors belonging to the public policy environment. Especially since public policy is the second most important stakeholder - after the private sector - in the total number of actors that have been identified.	Difficulties on the part of the Latitud foundation's RRP to initiate communication with public policy actors and effectively communicate the processes and results of the applied research carried out in the laboratory and how this can contribute to the design and implementation of public policy instruments for the sustainable development of the Uruguayan rice sector.	Actor 1, Actor 2, Actor 3, Actor 4, Actor 5, Actor 6
In relation to incipient PIs, particularly those actors belonging to the basic research environment, it is recommended that efforts be redirected towards dynamizing the evolution of these trajectories in terms of typology, depth and bidirectionality.	Difficulties in deepening the link. In the research environment (basic and applied), it is relatively easy to establish type I PIs, especially as these are those that occur on a personal level. In other words, researchers contact each other. However, the difficulty lies in the strategic dimension of this linkage and, above all, in the evolution of the trajectory of this linkage in the direction of fulfilling the final impacts identified by the programme theory.	Actor 7, Actor 8, Actor 9

Fig. 14. - Cross-referencing recommendations and conditioning elements for the identification of the reduced set of actors. Case study: RRP.

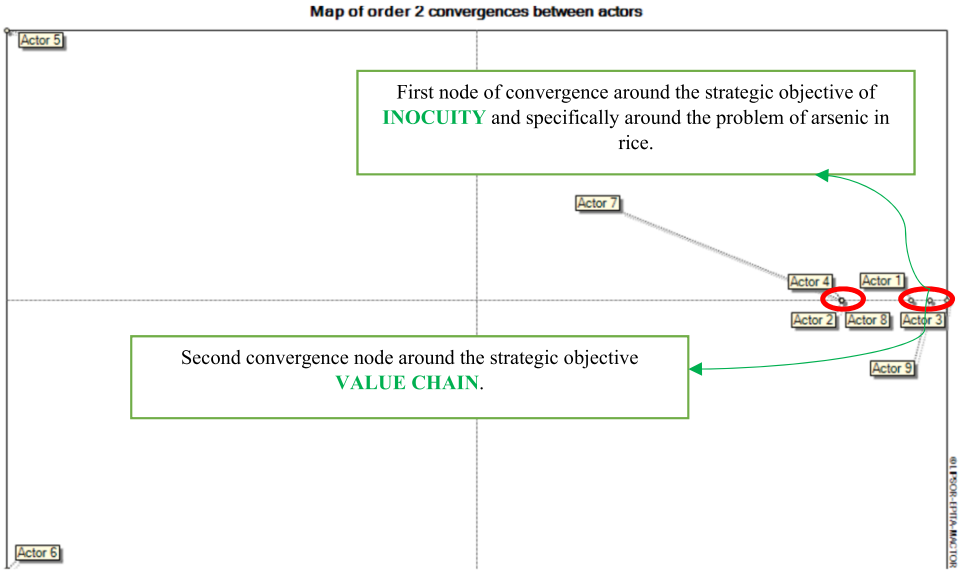


Fig. 15. - Output of the MACTOR software. Order 2 map of convergences between actors according to their hierarchy of objectives. Case study: RRP.

4. Discussion and conclusions from the process of applying the methodology to the case study

4.1. Discussion of contributions

The AE application process and the achieved results offer a strategic perspective to the evaluated program leader. It is a formative process that enables strategic learning through impact analysis, focusing on the mechanisms required to achieve impact and their connection to the program's environment.

As a theory-driven evaluation, it provides a comprehensive understanding of the program. Step 2 focuses on validating the theory within the context of the actual evaluation (C. H. Weiss, 1972). Similar to the *Impact Pathway* (Kuby, 1999) and *Developmental Evaluation* (Michael Quinn Patton, 2016) methodologies. These methodologies identify program theory as an impact pathway, particularly relevant in the field of science, technology, and innovation (STI) evaluation. One criticism of these approaches is the potential lack of a relevant and defensible theory of program logic, which can impact the success of the evaluation. In our case study, the theory was reviewed and validated by the lead researcher.

Similar to objective-based studies (Tyler, 1942), the approach is easily applicable to projects with clear and justifiable objectives. However, like other decision/accountability approaches, it may offer a

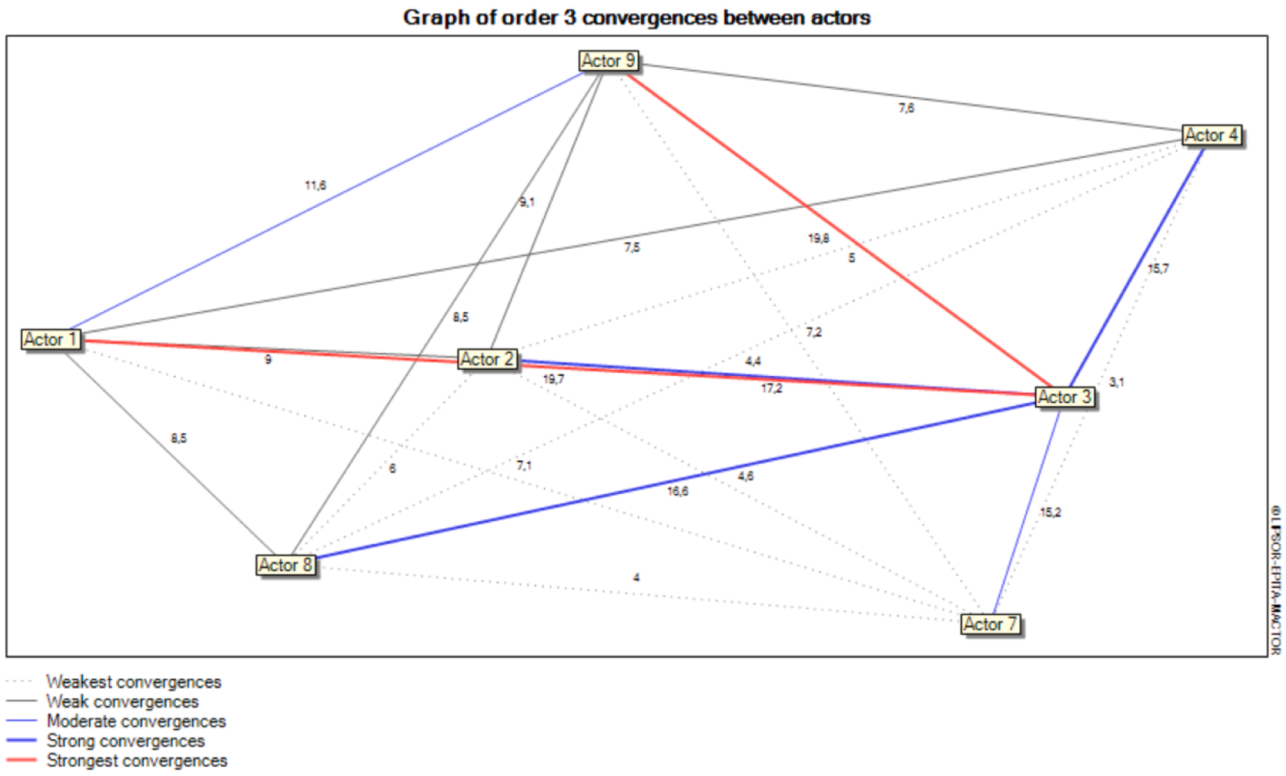


Fig. 16. Convergences between actors considering the hierarchy of objectives and the power relations between the actors themselves. Case study: RRP.

All the actors interviewed position themselves favourably in relation to the three specific strategic objectives of Latitud foundation's RRP. The only exception is actor 5, who is neutral, although in the course of the interview he identifies possibilities for convergence.

The differences between the actors are to be found in the hierarchies that the interviewees assign to these objectives, and this has to do with the (more or less indirect) compatibility they see with the strategic objectives of the institution in which they work. There are also differences in the power relations between the actors, with actors 2 and 3 standing out as the most influential or willing to coordinate, while actor 4 is one of the most isolated.

When analysed by objectives, two clear nodes of convergence can be observed. On the one hand, the node around the INOCUITY objective, which is ranked first by actors 3, 9 and 1. On the other hand, the node around the CHAIN VALUE objective, which is ranked by actors 2, 8, 4 and 7.

Fig. 17. - Synthesis of the analysis of the results obtained by the MACTOR method. Case study: RRP.

limited set of summary indicators. Nevertheless, this limitation is alleviated in step 3, which provides valuable information for decision-making and improved program management.

Emphasizing the importance of establishing productive interactions (PIs) with diverse actors, step 3 identifies a broad set of actors to achieve the committed impacts. This aligns with the purpose of the *Flows of Knowledge Framework* (Meagher, Lyall, & Nutley, 2008), recognizing that science can generate impact through productive interactions with society (Spaapen et al., 2012).

Decision/accountability approaches, referencing authors such as Alkin (1969) and Webster (1995), contribute to program staff's decision-making and emphasize improvement while serving as a basis for accountability. In our case study, the coverage indicator of 58 % demonstrates the program's significant effort to establish linkages during its implementation cycle. However, a limitation of these approaches, shared with our contribution, is the potential bias introduced through collaboration between the evaluator and program managers. External meta-evaluation, proposed by Stufflebeam (2001), can counteract this bias and ensure a balanced assessment.

The application of AE process facilitates the dynamization of PIs and the initiation of new ones. By applying the actors' game, convergence nodes around the program's strategic objectives can be identified. In our case study, we observed that certain PIs initiated by the program leader with researchers in the applied research environment have stagnated. Through the MACTOR method, we identified an actor (actor 9) who highly prioritized the program's strategic objective of INOCTUITY. This coincided with the positioning of two public policy actors (actor 1 and actor 3) with whom the program leader had not yet established a link. Identifying this convergence node allowed the program leader to deepen the connection with actor 9 and invigorate the trajectory of this PI.

AE shares characteristics with client-centered studies (or responsive evaluation) (Stake, 1983) by exploring the client's main interests and seeking relevant information. It examines the program's relationship, background, process, and outcomes, utilizing qualitative methods and triangulating results from different sources. A weakness of this approach is its potential lack of external credibility due to the close collaboration

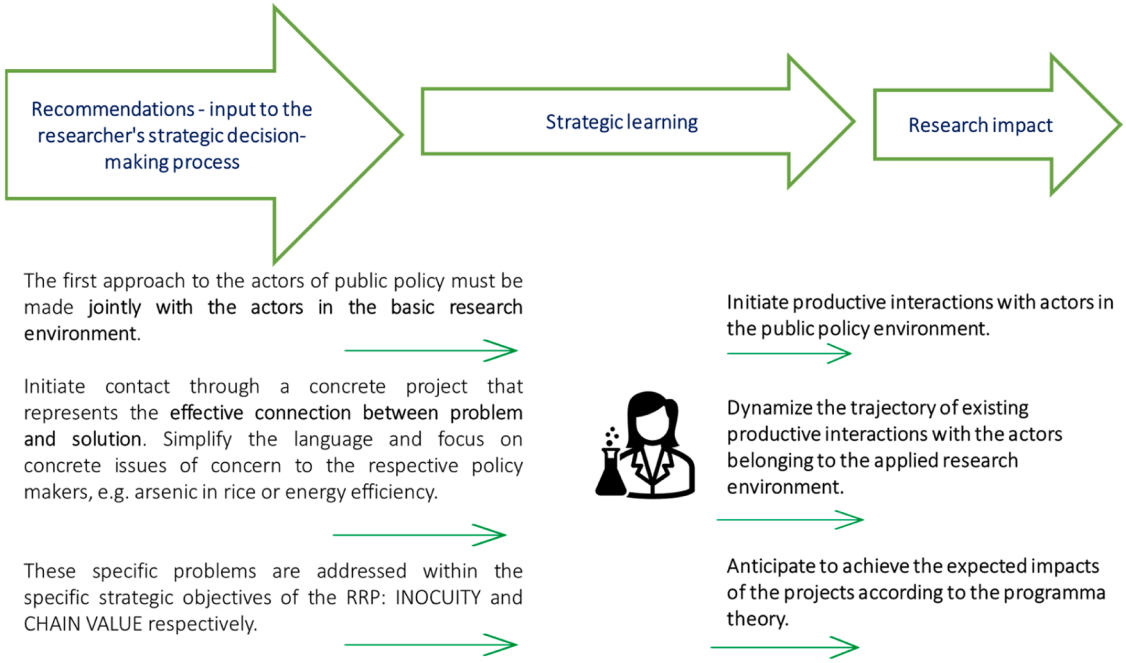


Fig. 18. - Synthesis of the analysis of the results obtained by the MACTOR method. Case study: RRP.

between evaluators and clients. External meta-evaluation can help mitigate this bias.

Quantitative indicators and qualitative analysis are combined in AE. Descriptive synthesis indicators (step 5) support the analysis of a snapshot of the established PI scheme at a specific point in time. Quantitative information is collected through semi-structured interviews using the MACTOR method. However, in-depth qualitative probing (step 8) complements this information, allowing for the formulation of recommendations (step 10). The mixed methods approach aims to provide guidance for improving programs as they evolve (Guba & Lincoln, 1989).

The anticipatory nature of this approach shares similarities with methodologies such as *Constructive Technology* Assessment (Guston & Sarewitz, 2002). This form of assessment is flexible in nature and allows for an intermediate positioning between ex-ante and ex-post approaches. In our case study we have applied it in the middle of the programme life cycle. This differs from ex-post approaches such as the *Advanced Technology Program Impact Assessment* (Ruegg & Feller, 2003), *Public Value Mapping* (Bozeman, 2003), *Payback Framework* (Buxton, 2011), *Contribution Mapping* (Kok & Schuit, 2012), *Socio-Economic Analysis of the Impact of Public Agricultural Research (*ASIRPA) (P. B. Joly et al., 2015) or *Research Contribution Framework* (Morton, 2015).

Through this approach, we obtain intermediate results that show the progress and improvement in achieving the committed impacts of the program, addressing the problem of temporality. This approach is also compared to the utilization-focused evaluation approach (M. Patton, 2008), where decisions about the evaluation study are made collaboratively with priority users, focusing on the intended uses of the evaluation. Utilization-focused evaluation allows for the combination of different techniques and methodologies, but its limitation lies in the potential influence of user groups with conflicts of interest.

The application of AE can contribute to alleviating the pressure researchers face to fulfil their scientific and societal mission at the same time. In our case study, the lead researcher has published in peer-reviewed journals, and with these metrics she can account for the fulfilment of her scientific mission. By applying AE, the researcher can also account for her social mission. By implementing 50 % of the programme, she has established links with 58 % of the key actors to achieve impact. She can also demonstrate that she has identified a specific strategy to unblock the trajectory of those PIs that have stalled and initiate the pending ones.

This clear intention to maximise expected impacts is a feature that coincides with the Rapid Outcome Mapping Approach (ROMA) (Young & Mendizabal, 2009) just as the focus on strategic research objectives has points in common with the prospective and adaptive evaluation of societal challenges *(PESCA)* approach (Weber & Polt, 2014) to mission-oriented policy evaluation.

Finally, we understand AE to be a tool that helps to manage research projects or programmes to achieve the expected impacts and in this sense it could be added to the Real Time Evaluation (RTE) family of tools (P.-B. Joly, Matt, & Robinson, 2019).

4.2. *Limitations and possibilities for future research*

In terms of opportunities for improvement, the actor identification process - broad set-in step 3 and narrow set-in step 7 - presents an opportunity. In our case study, we identified the broad set based on the programme logic model and by answering the question: which actors should be involved to achieve this component of the model? This allowed us to arrive at a first broad set of 36 actors. Then, in the presentation of results and recommendations to the programme team, 2 more actors were identified. The final set was therefore 38 actors. As for the reduced set, 9 actors were identified. In both cases, publicly available external information was triangulated. As mentioned in the development of step 3, official websites of the reference organizations, public information on project launches and even official press releases from authorities of the institutions identified as key. However, it is still a subjective process that may omit actors and therefore it is a process that can be improved in future applications.

The level of commitment of the actors could condition the application of the methodology and lack of commitment is also information for the evaluation. In our case study, 8 of the 9 semi-structured in-depth interviews were conducted. In 7 cases out of the 8 interviews a climate of trust and reflection was achieved, with an emphasis on learning opportunities. Only in one interview it was not possible to overcome the defensive attitude of the interviewee who perceived the interview as an evaluation of his work. In any case, it is all valuable and useful information to improve the understanding of the functioning of the system of

interest. However, a line for future research is identified here in terms of improving the process of identification and communication with stakeholders to strengthen their engagement or minimize dependence on it in terms of implementation success.

The criteria for classifying PIs by depth and bidirectionality could be improved and made more sophisticated. We understand that these criteria can be considered basic and that they allow for other levels of classification. In this sense, a clear line to deepen and give continuity to this work is made explicit. As an example, the bidirectionality that accounts for this co-creation process and ultimately for the change in the actor's way of doing things can be estimated with traditional impact indicators from scientometrics in the case of STI.

Also, in relation to characterising a detailed description of the PI, it could be the task of future research to use Social Network Analysis (SNA) to complement the MACTOR analysis by providing information on the relationships between stakeholders. For example, SNA can be used to identify: i) *central stakeholders*; these are stakeholders with a high degree of connectivity within the network. Central stakeholders are often influential because they have access to a large number of other stakeholders, ii) *bridging actors*; these are stakeholders who connect different stakeholder groups within the network. Bridging stakeholders can play a key role in facilitating communication and cooperation between different groups and, iii) *isolated stakeholders* are stakeholders who are not connected to other stakeholders in the network. Isolated stakeholders may be less influential because they have limited access to resources and information.

By combining the knowledge of MACTOR and SNA, decision-makers can better understand the stakeholder landscape.

Gender and diversity analyses could also be included in this characterisation of the PIs. For example, in the synthesis indicators in step 5, gender distribution indicators could be added for both the total number of identified stakeholders and existing and pending PIs.

This could facilitate the formulation of recommendations (step 6) in terms of increasing inclusive participation, identifying equity gaps between various groups, and identifying inclusive strategies to address these gaps in terms of power, influence and access to resources.

Although it is not the purpose of this paper, it is noted that future research may question the fundamental assumptions on which the EA has been formulated, for example whether PIs are transparent proxies for impact.

When we refer to the field of STI, we are aware that it is a broad field. STI indicators tend to have a broader meaning that clearly includes innovation activities. Scientometrics is often limited to science metrics. The purpose of using them as synonyms is to convey that, although science and innovation indicators are different, they cover overlapping spaces and face similar challenges (Ràfols, 2018). Nevertheless, we understand that it is a line of work to apply EA in programs and projects in diverse contexts to continue testing its usefulness as well as analyzing limitations and opportunities for improvement, which could change depending on the context of application.

5. Lessons learned

The application of this contribution has made it possible to demonstrate its usefulness for a particular case as well as to exemplify the contributions within the framework of the discussion presented in the previous section.

The applicability of the approach in different contexts will depend on the methodological needs identified by the evaluator. The previous section has mentioned the limitations but also the possibilities for future extensions in relation to the process of stakeholder identification, communication with stakeholders to strengthen their engagement, deepening the characterisation of the PIs, among others.

Other lessons learned are mentioned below, which could be useful for the evaluator considering replicating this methodology to be able to weigh up the difficulties, obstacles, and contributions of the application of the EA with as much information as possible.

Constant, focused, and reflective communication between the evaluator and the program leader is of utmost importance. The first two phases of implementation involved 8 weeks of work with weekly meetings lasting between 60 and 90 minutes. The success of the implementation and the establishment of a climate of reflection and trust depended on consistent communication, organized with clear frequency, and focused on short periods of time.

This indicator holds potential for future applications. The complete methodology required 21 weeks of work, which accounted for 9 % of the total program implementation time. It would be useful to know this indicator for other applications of EA in other contexts to estimate under what conditions the application of the methodology is feasible.

It is also worth mentioning that this application did not count on resources additional to those allocated to research within the framework of a PhD project.

AE allows working with different levels of aggregation. In our case study we worked with aggregation at the research programme level. The evaluation process was centred on the RRP lead researcher, who centralised the PI information of the entire research team under her responsibility. However, there was also the flexibility to integrate the entire programme team in some of the weekly meetings of the first stage to validate and favour the coverage and quality of the information systematised by the evaluator.

In the evaluation process, every piece of information is valuable. Unfortunately, the public policy actor, whom we intended to interview, did not respond to three emails and three phone calls. Contact was made with the actor two months after the application period had ended. This situation highlights the difficulty in obtaining timely responses, not only for the evaluator, who operates in the realm between science and politics and sought the interview, but also for the researcher.

The case study presented in this article belongs to the applied research environment. The application of the methodology yielded valuable results and recommendations for the researcher, her team, and the research center as a whole. However, we believe it is essential to continue testing the methodology in different science, technology, and innovation environments to improve and refine the proposal, as well as explore other fields of intervention.

In summary, the analysis of contributions, limitations, and possible extensions, as well as lessons learned, is intended to serve as an input for other evaluators to reflect on the suitability of the approach for their own purposes.

Funding

Funding for open access charge: CRUE-Universitat Politècnica de València.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

CRediT authorship contribution statement

Mariangel Pacheco - Troisi: Conceptualization, Data curation, Research, Methodology, Writing - original draft, Formal analysis, Software, Visualization. **Mónica García-Melón**: Research, Supervision, Writing - revision and editing, Methodology. **Fernando Jiménez-Sáez**: Conceptualization, Research, Supervision, Writing - review and editing, Methodology.

References

Ackoff, R. L. (1973). In E. Les (Ed.), *Méthodes de planification dans l'entreprise*. Paris: D'organisation.

Alkin, M. C. (1969). Evaluation theory development. *Evaluation Comment*. Retrieved from https://files.eric.ed.gov/fulltext/ED057208.pdf#page=41.
Alkin, M. C. (2011). *Evaluation Essentials* (Vol. 4). The Guilford Press.
Bozeman, B. (2003). *Public Value Mapping of Science Outcomes: Theory and Method: A Monograph of the Public Value Mapping Project of the Center for Science, Policy and Outcomes*. 3–48. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.105.6459&rep=rep1&type=pdf%0Ahttp://cord.asu.edu/wp-content/uploads/2015/02/Bozeman-B.-2003.pdf.
Buxton, M. (2011). The payback of "payback": Challenges in assessing research impact. *Research Evaluation, 20*(3), 259–260. https://doi.org/10.3152/095820211X13118583635837
Castro-Martínez, E., Olmos-Peñuela, J., & Fernandez-de-Lucio, I. (2016). La Vinculación Ciencia-Sociedad: Estereotipos y Nuevos Enfoques. *Journal of Technology Management and Innovation, 11*(2), 121–129. https://doi.org/10.4067/S0718-27242016000200012
Chen, H. T. (1990). *Theory-Driven Evaluations*. California SAGE.
Cozzens, S., & Snoek, M. (2010). Knowledge for policy. Contributing to the measurement of social, health, and environmental benefits. *Social Service Review, 55*(4), 525–543. https://doi.org/10.1086/643957
D'Este, P., Ramos-Vielba, I., Woolley, R., & Amara, N. (2018). How do researchers generate scientific and societal impacts? toward an analytical and operational framework. *Science and Public Policy, 45*(6), 752–763. https://doi.org/10.1093/SCIPOL/SCY023
Damşa, C. I. (2014). The multi-layered nature of small-group learning: Productive interactions in object-oriented collaboration. *International Journal of Computer-Supported Collaborative Learning, 9*(3), 247–281. https://doi.org/10.1007/s11412-014-9193-8
De Rijcke, S., Wouters, P. F., Rushforth, A. D., Franssen, T. P., & Hammarfelt, B. (2016). Evaluation practices and effects of indicator use-a literature review. *Research Evaluation, 25*(2), 161–169. https://doi.org/10.1093/reseval/rvv038
Durance, P. (2011). La Prospectiva Estratégica. *Gestión. Eno¯sis El Tercer Milenio, 5*(10), 61–75.
Godet, M. (1991). Actors' moves and strategies: The mactor method. An air transport case study. *Futures, 23*(6), 605–622. https://doi.org/10.1016/0016-3287(91)90082-D
Godet, M., & Durance, P. (2007). Prospectiva Estratégica: problemas y métodos. *Cuadernos Lipsor*.
Guba, E., & Lincoln, Y. (1989). *Fourth Generation Evaluation*. Retrieved from http://sgpwe.izt.uam.mx/pages/egt/Cursos/MetodoLicIII/7_Guba_Lincoln_Paradigmas.pdf.
Guston, D. H., & Sarewitz, D. (2002). Real-time technology assessment. *Technology in Society, 24*(1–2), 93–109. https://doi.org/10.1016/S0160-791X(01)00047-1
Hessels, L. K., van Lente, H., Grin, J., & Smits, R. E. H. M. (2011). Changing Struggles for Relevance in Eight Fields of Natural Science. *Industry and Higher Education, 25*(5), 347–357. https://doi.org/10.5367/ihe.2011.0056
Joly, P.-B., Matt, M., & Robinson, D. K. R. (2019). Research Impact Assessment: from ex post to real-time assessment. *Fteval Journal for Research and Technology Policy Evaluation, 47*(47), 35–40. https://doi.org/10.22163/fteval.2019.326
Joly, P. B., Gaunand, A., Colinet, L., Larédo, P., Lemarié, S., & Matt, M. (2015). ASIRPA: A comprehensive theory-based approach to assessing the societal impacts of a research organization. *Research Evaluation, 24*(4), 440–453. https://doi.org/10.1093/reseval/rvv015
Kok, M., & Schuit, A. (2012). Contribution mapping: A method for mapping the contribution of research to enhance its impact. *Health Research Policy and Systems, 10*, 1–16. Retrieved from http://www.health-policy-systems.com/content/10/1/21%5Cnhttp://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed10&NEWS=N&AN=2012495258.
Krange, I. (2007). Students' conceptual practices in science education: Productive disciplinary interactions in a participation trajectory. *Cultural Studies of Science Education, 2*(1), 171–203. https://doi.org/10.1007/s11422-006-9040-y
Kuby, T. (1999). *Innovation Islem- a Social Process, CIAT Workshop, Costa Rica, 4300. September 1999*. (August).
Ligero Lasa, J. A. (2015). *Tres Megyeétodos Dèlelőtt evaluaciòN Dèlelőtt Programas York Servicios Juicios finales sumativos, teoría del Cambio York evaluaciòN orientada a los actores implicados*.
Maassen van den Brink, H., de haas, M., van den Heuvel, J., Spaapen, J., Elsen, M., Westenbrink, R., & … van Drooge, L. (2010). Evaluating the societal relevance of academic research: A guide. *Evaluating Research in Context*, 1–24. https://doi.org/10.1007/978-94-007-5863-6_4
Manrique, S., Wroblewska, M. N., & Good, B. (2019). Rethinking research impact assessment: A multidimensional approach. *48*(48), 159–175. https://doi.org/10.22163/fteval.2019.385
Martin, B. R. (2010). The origins of the concept of "foresight" in science and technology: An insider's perspective. *Technological Forecasting and Social Change, 77*(9), 1438–1447. https://doi.org/10.1016/j.techfore.2010.06.009
Matt, M., Gaunand, A., Joly, P. B., & Colinet, L. (2017). Opening the black box of impact – Ideal-type impact pathways in a public agricultural research organization. *Research Policy, 46*(1), 207–218. https://doi.org/10.1016/j.respol.2016.09.016
Meagher, L., Lyall, C., & Nutley, S. (2008). Flows of knowledge, expertise and influence: A method for assessing policy and practice impacts from social science research. *Research Evaluation, 17*(3), 163–173. https://doi.org/10.3152/095820208X331720
Melloni, E., Pesce, F., & Vasilescu, C. (2016). Are social mechanisms usable and useful in evaluation research? *Evaluation, 22*(2), 209–227. https://doi.org/10.1177/1356389016643900
Molas-Gallart, J., & Tang, P. (2011). Tracing "productive interactions" to identify social impacts: An example from the social sciences. *Research Evaluation, 20*(3), 219–226. https://doi.org/10.3152/095820211X12941371876706
Morton, S. (2015). Progressing research impact assessment: A "contributions" approach. *Research Evaluation, 24*(4), 405–419. https://doi.org/10.1093/reseval/rvv016
Patton, M. (2008). *Utilization-focused evaluation*.
Patton, M. Q. (2002). Feminist, yes, but is it evaluation? *New Directions for Evaluation*.
Patton, Michael Quinn (2016). What is essential in developmental evaluation? on integrity, fidelity, adultery, abstinence, impotence, long-term commitment, integrity, and sensitivity in implementing evaluation models. *American Journal of Evaluation, 37*(2), 250–265. https://doi.org/10.1177/1098214015626295
Ràfols, I. (2018). S&T Indicators 'In the Wild': Contextualisation and Participation for Responsible Metrics (October 2017) *SSRN Electronic Journal*, 1–29. https://doi.org/10.2139/ssrn.3243134.
Ruegg, R., & Feller, I. (2003). *A Toolkit for Evaluating Public RD Investment*, 1–506.
Sivertsen, G., & Meijer, I. (2020). Normal versus extraordinary societal impact: how to understand, evaluate, and improve research activities in their relations to society? *Research Evaluation, 29*(1), 66–70. https://doi.org/10.1093/reseval/rvz032
Smit, J. P., & Hessels, L. K. (2021). The production of scientific and societal value in research evaluation: A review of societal impact assessment methods. *Research Evaluation, 30*(3), 323–335. https://doi.org/10.1093/reseval/rvab002
Spaapen, J., Shinn, T., & Marcovich, A. (2012). *SIAMPI Final Report Executive summary*, 1–36.
Spaapen, J., & van Drooge, L. (2011). Introducing "productive interactions" in social impact assessment. *Research Evaluation, 20*(3), 211–218. https://doi.org/10.3152/095820211X12941371876742
Stake, R. E. (1983). Program evaluation, particularly responsive evaluation. *Evaluation Models. (IV Social Agenda-Directed (Advocacy) Models)*, 342–362.
Stufflebeam, D. (2001). Evaluation models. *New Directions in Program Evaluation, 2001*(89), 7–98. https://doi.org/10.1002/ev.3
Tyler, R. W. (1942). General statement on evaluation. *Journal of Educational Research, 35*(7), 492–501. https://doi.org/10.1080/00220671.1942.10881106
Ward Hood, D., & Cassaro, D. A. (2002). Feminist evaluation and the inclusion of difference. *New Directions for Evaluation, 2002*(96), 27–40. https://doi.org/10.1002/ev.64
Weber, K. M., & Polt, W. (2014). *Assessing Mission-orientated R D programs: combining Foresight and evaluation*, (39), 5–10.
Webster, W. J. (1995). The connection between personnel evaluation and school evaluation. *Studies in Educational Evaluation, 1970*.
Weiss, C. H. (1972). *Evaluation*. Englewood Cliffs, NJ: Prentice Hall.
Weiss, Carol H. (1993). Where Politics and Evaluation Research Meet. *American Journal of Evaluation, 14*(1), 93–106. https://doi.org/10.1177/109821409301400119
Wolf, B., Lindenthal, T., Szerencsits, M., Holbrook, J. B., & Heß, J. (2013). Evaluating research beyond scientific impact: How to include criteria for productive interactions and impact on practice and society. *Gaia, 22*(2), 104–114. https://doi.org/10.14512/gaia.22.2.9
Young, J., & Mendizabal, E. (2009). Helping researchers become policy entrepreneurs. *Overseas Development Institute, 53*(September), 1–4. https://doi.org/10.1063/1.1671652

Mariangel Pacheco – Troisi: Predoctoral researcher at the Polytechnic University of Valencia - PhD (c), Master in Evaluation from the Complutense University of Madrid. Associate Professor at the Technological University of Uruguay and consultant specialized in the design and implementation of evaluations in the field of science, technology and innovation.

Mónica García-Melón: Full Professor in Industrial Engineering at the Polytechnic University of Valencia, Specialist in multicriteria decision analysis in participatory environments and indicators to monitor Responsible Research and Innovation in organizations.

Fernando Jiménez-Sáez: Tenured lecturer in Industrial Engineering at the Polytechnic University of Valencia, specialised in Economics of Technical Change and S&T policy evaluation.