Use of evaluation results in S&T and R&D: what does literature tell us?

DANIELA MACIEL PINTO[I]
ADRIANA BIN[II]
http://dx.doi.org/10.22347/2175-2753v18i58.5082

Abstract
This paper explores the use of evaluation results in Science and Technology (S&T) programs and policies, and Research and Development (R&D) through a systematic literature review. It investigates which knowledge fields have used evaluation results from 1970 to 2020, identifying four types of uses, along with theoretical, empirical, or mixed approaches, methods used, and document similarities. Using the PRISMA protocol and unsupervised data analysis, the study examines 62 studies, revealing a prevalence of studies in S&T policies and programs, with Social Sciences and Agricultural Sciences being the most common fields. An evolution from an instrumental approach focused on resource allocation to a more conceptual use related to the reflexivity of S&T and R&D activities is noted. The research primarily originates from the Northern Hemisphere, highlighting the need to systematize evaluations as valuable information sources.
Keywords: Research evaluation; Use of evaluations; Policy and program evaluation; Science and Technology; Research and Development; S&T Management; R&D Management.

Submetido em: 20/09/2024
Aprovado em: 02/03/2026

[I] Universidade Estadual de Campinas (UNICAMP), Campinas (SP), Brasil; https://orcid.org/0000-0003-0539-622X; e-mail: daniela.macielp@gmail.com.
[II] Universidade Estadual de Campinas (UNICAMP), Campinas (SP), Brasil; https://orcid.org/0000-0001-9774-2753; e-mail: adribin@unicamp.br.

Uso de resultados de avaliações em C&T e P&D: o que diz a literatura?

Resumo
Este artigo explora o uso de resultados de avaliações em programas e políticas de Ciência e Tecnologia (C&T) e em Pesquisa e Desenvolvimento (P&D), por meio de uma revisão sistemática da literatura. Investiga quais campos do conhecimento utilizaram resultados de avaliações entre 1970 e 2020, identificando quatro tipos de uso, bem como as abordagens teóricas, empíricas ou mistas, métodos utilizados e similaridades entre os documentos. Utilizando o protocolo PRISMA e análise de dados não supervisionada, o estudo examina 62 trabalhos, revelando a prevalência de estudos sobre políticas e programas de C&T, com destaque para as Ciências Sociais e Ciências Agrárias. Observa-se uma evolução de um uso instrumental, focado na alocação de recursos, para um uso mais conceitual, relacionado à reflexividade das atividades de C&T e P&D. A pesquisa é majoritariamente oriunda do Hemisfério Norte, destacando a necessidade de sistematizar avaliações como fontes valiosas de informação.
Palavras-chave: Avaliação da pesquisa; Uso de avaliações; Avaliação de políticas e programas; Ciência e Tecnologia; Pesquisa e Desenvolvimento; Gestão de C&T; Gestão de P&D.

Uso de resultados de evaluaciones en C&T e I+D: ¿qué nos dice la literatura?

Resumen
Este artículo explora el uso de los resultados de evaluaciones en programas y políticas de Ciencia y Tecnología (C&T) y en Investigación y Desarrollo (I+D), a través de una revisión sistemática de la literatura. Investiga qué campos del conocimiento han utilizado resultados de evaluaciones entre 1970 y 2020, identificando cuatro tipos de uso, junto con enfoques teóricos, empíricos o mixtos, métodos empleados y similitudes entre los documentos. Utilizando el protocolo PRISMA y análisis de datos no supervisado, el estudio examina 62 trabajos, revelando una prevalencia de estudios sobre políticas y programas de C&T, destacándose las Ciencias Sociales y las Ciencias Agrícolas. Se observa una evolución desde un uso instrumental, centrado en la asignación de recursos, hacia un uso más conceptual, relacionado con la reflexividad de las actividades de C&T e I+D. La investigación proviene principalmente del Hemisferio Norte, lo que resalta la necesidad de sistematizar las evaluaciones como fuentes valiosas de información.
Palabras clave: Evaluación de la investigación; Uso de evaluaciones; Evaluación de políticas y programas; Ciencia y Tecnología; Investigación y Desarrollo; Gestión de C&T; Gestión de I+D.

Introduction

In Science and Technology (S&T) and Research and Development (R&D), evaluative activity has been conducted since the 1960s, primarily to verify the relationship between investment and return (accountability) (Cruz Castro; Sanz Menéndez, 2006; Cruz-Castro; Sanz-Menéndez, 2005). As challenges have become more complex and demands have expanded, these evaluations are now seen as sources of strategic information regarding institutional learning and social, economic, environmental, and political impacts (Pinto. Bin; Rodrigues, 2025; Milzow *et al*., 2019).

By S&T and R&D evaluation, we understand the investigation of the effectiveness of policies, programs, projects, and activities related to the promotion and execution of systematic activities aimed at advancing knowledge and technological development, whether in the governmental sphere or within public and private organizations (research institutes, higher education institutions, companies, among others).

Thus, for the context of the present article and for analytical purposes, we will consider that S&T evaluation refers to the investigation of policies and programs, adopting a more macro perspective, while R&D evaluation focuses on research projects and activities, with a more micro perspective, centered on research organizations. At the same time, we understand that the use of evaluation results involves employing in different ways the findings that a formal evaluation produces in both contexts: S&T and R&D, for different perspectives.

In this vein, studies suggest that the effectiveness of S&T policies and programs evaluation and R&D evaluation depends on their influence on strategic and operational decisions (Bozeman; Sarewitz, 2011; Cozzens, 2012). Cozzens (2012) argues that in S&T, evaluations should be more than bureaucratic items, acting as catalysts for significant changes. This requires managers, especially those of public resources, to have greater responsibility in applying investments towards a socially responsive and impactful agenda, as expected today (Felt, 2018; Stilgoe; Guston, 2016).

Thus, beyond accountability, which is fundamental, it is essential to understand how the results of S&T policies and programs and R&D evaluations are used or can be applied to support the formulation and refinement of more effective policies and strategies as well as to support the management and adjustment of activities aimed at advancing knowledge and technological development (Bozeman; Sarewitz, 2011; Cruz Castro; Sanz Menéndez, 2006; Foray; Mowery; Nelson, 2012). However, in S&T and

R&D, the discussion about the use of evaluation results lacks practical and theoretical investigations (Lee, 2010; Van der Most, 2010).

In light of this, the present work aims to conduct a systematic review of the literature to highlight how the topic has been addressed in the context of S&T policy and program evaluation, as well as in R&D evaluation. Thus, the question sought to be answered was: "how is the use of evaluation results presented over time in the specialized literature, considering knowledge fields, types of uses, and methods adopted"? For this purpose, the PRISMA protocol and techniques of content analysis and unsupervised data analysis were employed, investigating a set of 62 technical-scientific publications from three databases. To the best of our knowledge, this is the first systematic review of the literature, structuring findings by periods to address the lack of debates on the subject in the evaluation literature of S&T and R&D. The findings are relevant for evaluators, researchers, R&D managers, and S&T policy managers.

Theoretical Framework

Although distinguishing evaluations in these fields is complex - since S&T policies often relate to R&D - this study proposes a separation based on focus, context, objectives, and actors (Table 1).

Table 1 - Main differences between evaluation of S&T policies and programs and R&D

Criterion	Programs and Public Policies	R&D Management
Focus	Evaluation of S&T policies and programs for broad objectives (macro vision)	Evaluation of specific R&D projects and activities, with a more micro perspective
Application Context	Governmental or Institutional	Organizational or Corporative
Objectives	Ensure that policies and programs achieve their goals and objectives in short and long terms	Improve the quality, efficiency, and impact of R&D actions
Actors	Governments, funding agencies, legislators, public policymakers	Researchers, R&D managers, directors of R&D organizations

Source: The authors (2025).

Use of Evaluation Results: A Reflection

Since the 1960s, the use of evaluation results has been a central topic in program and public policy evaluation, with Alkin and Christie (2004; 2023) organizing the topic in what they call “Branch of use”. Included in this branch, Stufflebeam (1968;

1983; 2002) and Wholey (2004) pioneered the idea that evaluation design influences how results are used. In this sense, Stufflebeam (1983; 2002) introduced the first model focused on use: the CIPP model[III], which promotes learning, improvement, and accountability. Wholey (2004), in turn, stressed "evaluability," considering feasibility and utility for decision-makers. On the other hand, Patton (2008; 2015) developed the theory of Utilization-Focused Evaluation (UFE), which emphasizes conducting evaluations with a clear orientation toward use, regardless of a specific intended user. He argued that evaluation results should be applicable across diverse contexts and stakeholders. Patton (2015) also identified the evaluator as a value agent (someone who actively facilitates use) a role later reframed as that of a knowledge broker, a strategic actor who supports policymaking (Donnelly *et al*., 2014; Olejniczak, 2017; Olejniczak; Raimondo; Kupiec, 2016).

Rejecting this idea, Alkin *et al*. (1979) argued that users should establish value systems to judge evaluation results and not evaluators. Cousins *et al*. (2009) supports this view, and proposed participatory approaches to connect results with users. In this line, Weiss (1972; 1998) viewed use as a natural consequence of evaluation, shaped by beliefs and contextual shifts. While she originally categorized use into three types (instrumental, conceptual, and persuasive/symbolic), subsequent authors[IV] have proposed a fourth (Table 2), influence/process, reflecting learning and behavioral change during the evaluation process.

Table 2 - Types of Use

Type of Use	Definition
Instrumental	Refers to the direct use of evaluation results to make concrete decisions and inform practical actions. The results are applied tangibly to improve the program or policy.
Conceptual	Involves using the results to enhance the conceptual and theoretical understanding of a problem or issue. The results help to expand knowledge about a specific subject and refine the conceptual foundation of the program or policy.
Persuasive/Symbolic	Refers to the use of evaluation results to shape opinions, legitimize decisions, or construct narratives that support political, organizational, or strategic goals—often without actual application of findings.
Influence/Process	Refers to learning and changes in behavior, attitudes, or organizational culture resulting from participation in the evaluation process, regardless of findings. Includes instrumental and conceptual outcomes that emerge through engagement rather than through use of final results.

Source: Weiss (1998) and Alkin and King (2016).

[III] The model considers the need to identify the previously established goals (C), the resources made available for implementing an action (I), the process adopted for the implementation of the action (P), and the products (P) generated from the implementation (Stufflebeam, 1983).

[IV] The category "influence/process" is adopted here based on Patton (2008; 2018) and Alkin and King (2016), who emphasize learning and behavioral change resulting from participation in the evaluation process.

In program and policy evaluation, reviews and empirical studies since the 1960s (Cousins; Leithwood, 1986; Johnson *et al*., 2009; Leviton; Hughes, 1981) have highlighted key factors influencing the use of results, such as evaluation quality, stakeholder engagement, and institutional capacity. These aspects led to the development of Evaluation Capacity Building (ECB), aimed at sustaining evaluation competencies in organizations (Bourgeois; Cousins, 2009), with models by Preskill and Boyle (2008; 2009) and Labin *et al*. (2012) guiding this process.

Use of Evaluation Results in S&T and R&D

In research evaluation, however, such foundations remain limited. Although the Branch of Use offers a useful basis for studies on the use of evaluation results in S&T and R&D, the topic is still underexplored and lacks theoretical and practical development (Cozzens; Melkers, 1997; Milzow *et al*., 2019; Van der Most, 2010). This limited discussion may partly reflect the field's emphasis on metric-based approaches, such as bibliometrics. Authors like Spaapen (2015) and Reed *et al*. (2021; 2022) criticize this dominance and call for methods better suited to research contexts.

Paradigms such as Responsible Research Assessment (RRA) and Responsible Research and Innovation (RRI), emerging in the late 2000s, respond to these critiques. RRI promotes agendas aligned with social demands, while RRA emphasizes effects on reality beyond scientific output (COARA, 2022; Felt, 2018; Schönbrodt *et al*., 2022; Von Schomberg, 2019). Both encourage reflexivity in science and evaluation, reinforcing the role of evaluation results in shaping inclusive and socially relevant S&T policies and R&D actions.

Institutions like the NSF and REF include the use of evaluation results in their frameworks, especially through reporting and communication. The NSF, for instance, provides guidance via the EvaluATE Hub (National Science Foundation, 2024). Although the UK's Research Excellence Framework (REF), active since 1986, does not explicitly define how results should be used (Northan, 2010; Research Excellence Framework, 2024), it promotes their use through performance-based funding. This encourages universities to apply evaluation results to improve research quality and internal practices (Morgan Jones; Grant; Rand, 2013; Morgan Jones; Manville; Chataway, 2017; Parks *et al*., 2019), highlighting what Parks *et al*. (2019) called the "A's of Evaluation": Accountability, Advocacy, Analysis, Allocation, Acclaim, and Adaptation.

Despite progress, studies examining how and why evaluation results are used in S&T and R&D remain scarce. Van der Most (2010) stands out, identifying 15 studies and mapping methods and users, though with limitations in study selection and context. This gap points to the need for further investigation. The literature of Branch of Use can support studies related to the use of evaluation results in S&T and R&D. It is important, however, to consider that the discussion of this theme, more oriented towards guidelines and activities aimed at advancing knowledge, is underexplored and lacks theoretical and practical investigations (Cozzens, 2012; Milzow *et al*., 2019; Van der Most, 2010).

Methodological Procedures

We applied the PRISMA protocol (Page *et al*., 2021a; 2021b) which follows a sequential method: (1) comprehensive search; (2) selection based on predefined criteria; (3) systematic data analysis; and (4) structured reporting to reduces bias and enhances review quality (Galvão; Pansani; Harrad, 2015; Moher *et al*., 2009).

Bibliographic Data Sources

Data were retrieved from Web of Science (WoS), Scopus, and Google Scholar (GS), databases selected for their relevance in scientific and technical literature (Harzing; Alakangas, 2016; Mongeon; Paul-Hus, 2016). While WoS and Scopus focus on journal articles, GS also indexes technical documents such as reports and institutional publications (Kousha; Thelwall, 2007). The search, unrestricted in time or geography, was conducted from August to October 2023, using three expressions with boolean operators, truncation (*), and quotes (“”) in titles and abstracts. A total of 252 studies were retrieved (see Table 3).

Table 3 - Search strategies in each database

Expression	Scopus	WoS	GS	Total
("use of evaluation" OR "utilization of evaluation" OR "using evaluation") AND ("research" OR "R&D" OR "RD" OR "research and devel*" OR "innov*")	28	20	107	155
("use of assess*" OR "utilization of assess*" OR "using assess*") AND ("research" OR "R&D" OR "RD" OR "research and devel*" OR "innov*")	11	5	74	90
"evaluation utilization" OR "assessment utilization" AND "research and development" OR "R&D" OR "RD" OR "Innovation" OR "Science" OR "Science and technology" OR "ST"	1	0	6	7
Amount	40	25	187	252

Source: The authors (2025).

It should be noted that the data were retrieved from the database sites, except for Google Scholar, which was obtained using the software Publish or Perish (Harzing, 2016).

Sample Selection

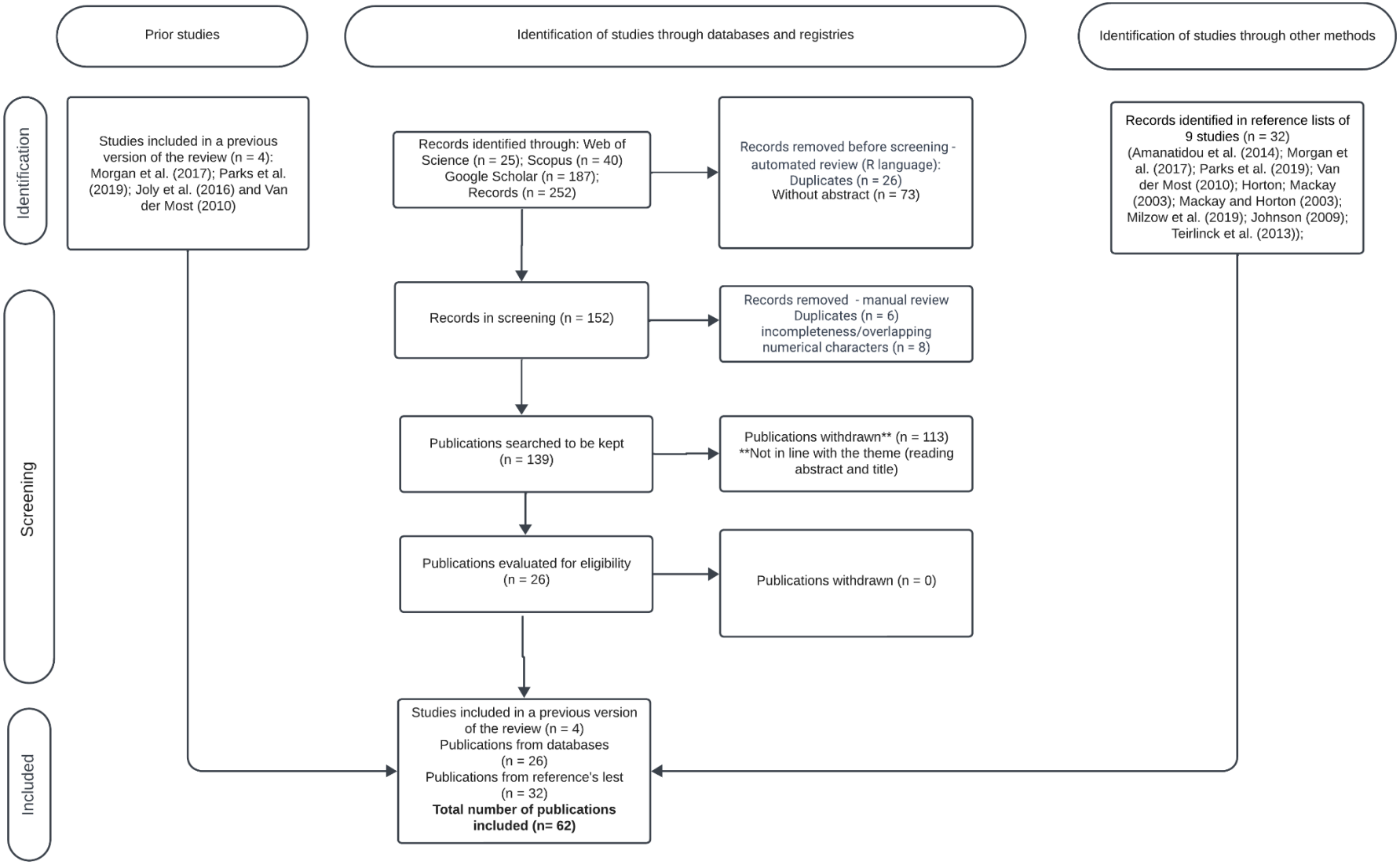
For the analysis and selection of the studies, a verification of title duplications and summary field completion was initially conducted. In this first step, the R language with the dplyr package was used, identifying 26 duplicates and 73 studies without a filled summary field, leaving 153 studies. A manual verification then followed, identifying an additional 6 duplicates and 8 studies with incomplete metadata or characters that made identification difficult. Thus, 139 articles were examined, considering two inclusion criteria: 1. relevance of the study, considering its direct linkage to S&T and/or R&D (n = 113 removed) and 2. studies available in full (n = 0 removed). Recognizing the intersection between S&T and R&D, the classification was based on Table 1 of the theoretical framework and:

- for S&T programs and policies: by explicit mention of government policies and/or programs with the presence of the words "science and technology" and its variations "S&T", "ST&I".
- in R&D: focused on the identification of research actions in organizational settings, such as public and private research centers.

Based on these criteria, 26 studies were selected for analysis. Additionally, four relevant studies[V] from the authors' personal archive, not retrieved through the database search, were included. Finally, an approach was adopted to ensure a more comprehensive review, through the analysis of the reference[VI] lists of the 26 selected studies. In this way, another 32 documents, originating from eight studies[VII], were incorporated into the analysis, totaling 62 selected studies, as recorded in Figure 1.

[V] Joly *et al.*, 2016; Morgan Jones; Grant; Rand, 2013; Morgan Jones; Manville; Chataway, 2017; Parks *et al*., 2019; Van der Most, 2010.
[VI] Manual analysis of each one of the 62 studies.
[VII] Amanatidou *et al*., 2014; Horton; Mackay, 2003; Mackay; Horton, 2003; Milzow *et al*., 2019; Morgan Jones; Grant; Rand, 2013; Morgan Jones; Manville; Chataway, 2017; Teirlinck *et al*., 2013; Van der Most, 2010.

Figure 1 - Study selection flowchart based on PRISMA recommendations



Source: The authors (2025).

Analysis of Studies

The 62 studies were analyzed considering: the identification of the knowledge fields to which they were related, the focus of evaluation (either S&T policy and program or R&D), the type of use identified, the methods adopted, considering the approach (either empirical or theoretical), the institutions/affiliations of the authors, and the countries in which the research took place. The knowledge fields were identified using the classification proposed by the OECD (Organization for Economic Cooperation and Development, 2002) into six major areas: 1. Natural Sciences; 2. Engineering and Technology; 3. Medical and Health Sciences; 4. Agricultural Sciences; 5. Social Sciences; and 6. Humanities. Thus, the studies were read searching for evidence related to the type of area, considering:

1. Natural Sciences: Studies directly related to physics, chemistry, biology, and geology, or which explore natural phenomena in a more isolated manner.
2. Engineering and Technology: Studies that investigated innovation and technological development in the industrial sector, primarily.
3. Medical and Health Sciences: Studies directly related to research in medicine and human health.
4. Agricultural Sciences: Studies related to research and development in the agricultural sector.
5. Social Sciences: Studies of the socioeconomic dynamics that shape the field of science and technology.
6. Humanities: Studies dedicated to the critical analysis of scientific and technological knowledge.

The focus of evaluation was primarily considered based on the differences presented in Table 1, from the theoretical framework of this study. The type of use was identified based on table 2, also presented in the theoretical framework. The identification of methods and approach was considered based on the structure presented in Table 4.

Table 4 - Type of approach and methods

Approach	Description
Empirical	Refers to studies that sought to investigate, in practice, how the use of evaluation results occurs in different contexts and environments of S&T and R&D. For example: Case studies; Interviews; Focus groups; etc.
Theoretical	Represents studies focused on theoretical discussion, knowledge advancement, or reflections about the theme. For example: Literature review; Theoretical discussion; Document analysis, etc.

Source: The authors (2025).

Institution and country identification was conducted via data manipulation of the "Affiliation" field from the databases, using R and the packages "tidyverse," "dplyr," and "tibble." To support the analysis of similarities among studies and organize the S&T and R&D evaluations, we applied two processes: tokenization and topic modeling. Tokenization, a Natural Language Processing (NLP) technique, identifies the most frequent words in a text corpus (Silge; Robinson, 2017), helping to highlight keywords common across studies. Topic modeling, a machine learning technique, identifies semantic patterns across texts to group documents by topic (Silge; Robinson, 2017). We applied this method to the full texts of 61 of the 62 selected studies. One exception was Lee (2010), for which the original file was unavailable; only metadata and indexed database information were used to include its contributions. Data and algorithms used in the study are accessible in the GitHub repository: https://bit.ly/469tClb.

Software Used

The software RStudio, version 2023.6.1.524, and the R language, version 4.3.1 (2023-06-16 ucrt), were adopted throughout the study. For organization, interpretation, and visualization of the data, LibreOffice Calc, version 7.5.5.2 (X86_64), and Tableau, professional edition version 2023.2.2, were complementary tools.

Findings

A diversification in document types and a progressive increase in scientific and technical production were identified in investigations of evaluation use in S&T policies and programs and R&D projects and activities. Approximately 75% of the studies relate to the use of evaluation results in S&T, while 25% pertain to R&D. According to Table 5, scientific articles and technical reports are the main types of documents in both types of investigation, representing 88% of the overall total.

Table 5 - Scientific production on the use of evaluation

Focus	Document Type	1970	1980	1990	2000	2010	2020	Total	%
S&T	Article	1	4	6	9	8	1	29	61,7
	Article Conference					2		2	4,3
	Mono					1	1	2	4,3
	Report			2	4	8		14	29,8
Total S&T Programs		**1**	**4**	**8**	**13**	**19**	**2**	**47**	**100**
R&D	Article			3	4	1		8	53,3
	Article Conference						1	1	6,7
	Editorial Material				1			1	6,7
	Mono					1		1	6,7
	Report				1	3		4	26,7
Total R&D		**0**	**0**	**3**	**6**	**5**	**1**	**15**	**100**
Overall Total		**1**	**4**	**11**	**19**	**24**	**3**	**62**	

Source: The authors (2025).

All knowledge fields proposed by the OECD were identified, with Social Sciences, Engineering and Technology, and Medical and Health Sciences being particularly prominent in discussions about the use of results in S&T policies and programs. In contrast, Agricultural Sciences, Social Sciences, and Medical Health Sciences were notable in R&D projects and activities (Table 6). 45% of the studies spanned multiple fields and were conducted across 13 different types of institutions from 17 countries. Among these institutions, 12 focused on the use of evaluation results in S&T policies and programs, compared to 5 that focused on R&D projects and activities, as illustrated in Figure 1. Notably, countries from the Northern Hemisphere were prominent, with the United Kingdom and the United States standing out.

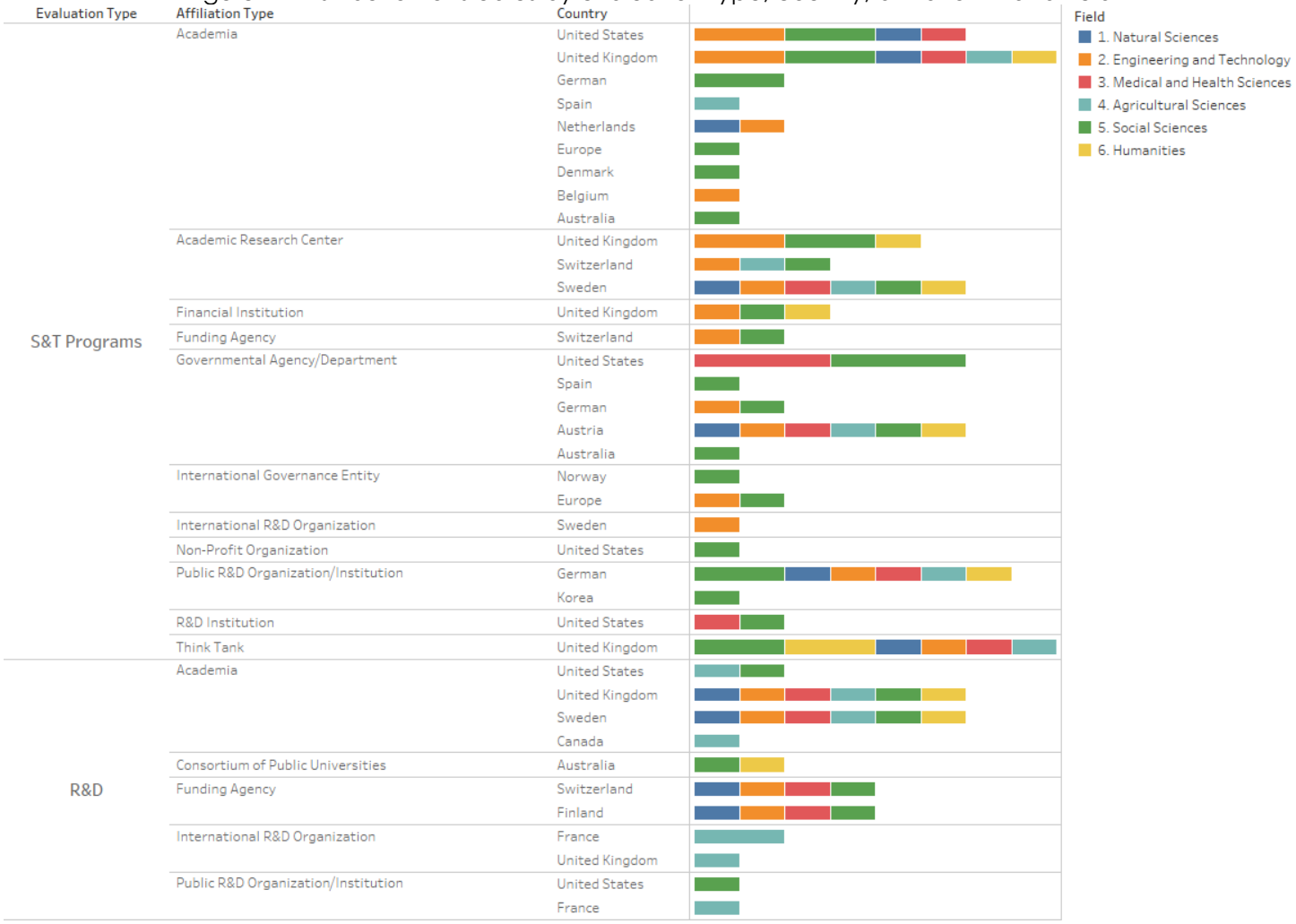
Various S&T policies and programs were investigated, highlighting European research funding programs, innovation initiatives for industry, and higher education programs aligned with the "Europe 2020" strategy with REF standing out. These encompassed a wide array of subjects, such as mental health programs, environmental protection, and efforts from the NSF program to advance mathematical research (MSP-RETA) and others. Regarding R&D, actions have involved organizations like the Consultative Group on International Agricultural Research (CGIAR), International Forestry Research Center (CIFOR), the National Institute of Agricultural Research (INRAE), France, the Swiss National Science Foundation (SNSF), and the Innovative Research Universities Australia (IRU) (Australia).

Table 6 - OECD areas identified by type of evaluation

Evaluation	Field	Authors	1970	1980	1990	2000	2010	2020	Total
ST&I	5. Social Sciences	Amanatidou (2014); Baker (2013); Balthasar; Rieder (2000); Barker (2007); Bayhan et al. (2011); Bornmann; Marx (2013); Boyer; Langbein (1991); Brofoss (1998); Brunet; Müller (2022); Carlsson et al. (1999); Chelimsky (1977); Chelimsky (2001); Flemming (1991); FTeval (2003); Gold et al. (2011); Gordon et al. (2017); Hall (1981); Hansson (2006); Johnson; Greenseid; Toal (2009); Kuhlmann (1998); Lee (2010); Macdonald (2006); Mackay K (1992); Manville et al. (2015); Meyer-Krahmer; Montigny (1989); Morgan et al. (2017); Most (2010); Northon (2010); Ramalingam (2011); Sanz Menendez (1997); Saunders et al. (2005); Shapira; Kuhlman (2003); Siegel;Tuckel (1985); Thakur (2007); Wollmann (2016)	1	3	7	10	12	1	34
	2. Engineering and Technology	Amanatidou (2014); Balthasar; Rieder (2000); Bayhan et al. (2011); Carlsson et al. (1999); Chelimsky (2001); Cozzens; Melkers (1997); Flemming (1991); FTeval (2003); Gök (2010); Gök; Edler (2012); Kuhlmann (1998); Meyer-Krahmer; Montigny (1989); Morgan et al. (2017); Most (2010); Roseland; Volkov (2011); Shapira; Kuhlman (2003); Smith (2008); Teirlinck et al. (2013); Ter Bogt; Scapens (2009); Thompson; Levitov (1983); Vakola (2000)	0	2	4	7	8	0	21
	3. Medical and Health Sciences	Amanatidou (2014); Boyer; Langbein (1991); Chelimsky (2001); FTeval (2003); Gold et al. (2011); Gordon et al. (2017); Kjesrud (2021); Kuhlmann (1998); Morgan et al. (2017); Most (2010); Shapira; Kuhlman (2003); Siegel;Tuckel (1985)	0	1	1	3	5	1	11
	6. Humanities	Amanatidou (2014); Chelimsky (2001); Flemming (1991); FTeval (2003); Kuhlmann (1998); Macdonald (2006); Manville et al. (2015); Morgan et al. (2017); Most (2010); Shapira; Kuhlman (2003); Ramalingam (2011); Parks et al.(2019)	0	0	2	4	5	0	11
	1. Natural Sciences	Amanatidou (2014); Chelimsky (2001); FTeval (2003); Kuhlmann (1998); Morgan et al. (2017); Most (2010); Roseland; Volkov (2011); Shapira; Kuhlman (2003); Ter Bogt; Scapens (2009); Thompson; Levitov (1983)	0	1	1	4	4	0	10
	4. Agricultural Sciences	Amanatidou (2014); Balthasar; Rieder (2000); Chelimsky (2001); Diez; Izquierdo; Malagon (2016); FTeval (2003); Kuhlmann (1998); Morgan et al. (2017); Most (2010); Shapira; Kuhlman (2003)	0	0	1	4	4	0	9
R&D	4. Agricultural Sciences	Bennett (2001); Grant; Hinrichs (2015); Hemlin; Sven (1996); Horton (1998); Horton; Mackay (2003); Joly et al. (2016); Mackay; Horton (2003); Patton; Horton (2009); Raitzer; Kelley (2008)	0	0	2	5	2	0	9
	5. Social Sciences	Grant; Hinrichs (2015); Hemlin; Sven (1996); Jordan; Zuckerman (2008); Luukkonen (1995); Milzow et al. (2019); Nelson (2005); Stephens-Chu (2022)	0	0	2	1	3	1	7
	3. Medical and Health Sciences	Grant; Hinrichs (2015); Hemlin; Sven (1996); Kamenetzky et al. (2016); Luukkonen (1995); Milzow et al. (2019)	0	0	2	0	3	0	5
	1. Natural Sciences	Grant; Hinrichs (2015); Hemlin; Sven (1996); Luukkonen (1995); Milzow et al. (2019)	0	0	2	0	2	0	4
	2. Engineering and Technology		0	0	2	0	2	0	4
	6. Humanities	Grant; Hinrichs (2015); Hemlin; Sven (1996); Nelson (2005); Parks et al.(2019)	0	0	1	0	2	0	3

Source: The authors (2025).

Figure 1 - Distribution of studies by evaluation type, country, affiliation[VIII] and field



Source: The authors (2025).

[VIII] The affiliation corresponds to the institution of the authors responsible for the evaluation and not necessarily to the institution of the subject being evaluated.

Evolution and Use of Evaluation Results in S&T and R&D

The topic of evaluation use emerged in the 1970s within the Social Sciences, in the context of S&T policy programs. Chelimsky (1977) provided a seminal analysis of how program managers in North American S&T programs used evaluation results, identifying both barriers and facilitating factors. In the 1980s, the topic expanded to Engineering and Technology, Medical and Health Sciences, and Natural Sciences, accompanied by methodological innovations. Hall (Hall, 1981), for instance, applied the Concerns-Based Adoption Model (CBAM) to explore the understanding and improvement of evaluation use in innovation. Thompson and Levitov (1983) developed a quantitative method to analyze evaluation use across American programs, while Siegel and Tuckel (1985) studied how well-constructed evaluations influence their application.

Throughout this early period (1970–1980), instrumental use of evaluation results dominated, as classified by Weiss (1998), appearing in 50 studies (80%) across all fields. Instrumental use refers to applying evaluation results directly in decision-making and action, often linked to funding agencies, budget governance, and strategic allocation of resources in R&D organizations and universities. The key users[IX] during this phase were policymakers and program managers, mainly.

The 1990s introduced new document types, like theses, dissertations, and technical reports, and saw the expansion of research on evaluation use into all OECD knowledge fields, with R&D becoming a more prominent focus. Studies examined how evaluations influenced public budgets and strategic research directions. In R&D, Hemlin (1996) and Luukkonen (1995) analyzed how evaluations guided internal funding decisions in Finland. Horton (1998), reflecting on the CGIAR context, showed evaluations as mediators between science and policy. Similarly, Mackay (1992), Kuhlmann (1998), Sanz Menéndez (1997), explored evaluations as tools for shaping S&T policies and guiding investment decisions.

During this period, conceptual use (identified in 32 studies (51%)) gained visibility. Conceptual use relates to using results to build theoretical understanding and refine frameworks. It was particularly linked to reflexivity in S&T and R&D, with Social Sciences, Engineering, and Agricultural Sciences as key fields. Early users included policymakers, program managers, and industry; from the 2000s, the focus shifted to

[IX]Adjusted based on a minimum of two studies.

universities, academic managers, funding bodies, healthcare, and agricultural institutions.

In the 2000s, there was a 44% increase in scientific publications, with Social Sciences maintaining a strong presence. Studies continued to analyze how evaluations affected funding decisions and began to treat evaluation as a strategic component of innovation and organizational learning. In R&D, impact evaluations became central to strategic management, especially in agriculture. Horton and Mackay (2003) and Mackay and Horton (2003) emphasized integrating evaluations into learning-based change strategies at CGIAR. Patton and Horton (2009) expanded this by proposing Utilization-Focused Evaluation (UFE) as a tool to support innovation. Raitzer and Kelley (2008) revealed occasional disconnects between evaluation findings and funding decisions.

Influence/process use, the third most cited type, appeared in 21 studies (33%) and was most frequent between the 1980s and 2010. Found in 77% of studies focused on S&T programs, it refers to evaluations shaping not only the evaluation process itself but also informing policy adjustments and strategic direction. It followed a user trajectory similar to instrumental use: from policymakers to inclusion of funders, R&D organizations, and academic managers.

In the 2010s, the number of articles declined by 38%, but technical publications rose by 37%. Engineering and Medical Sciences advanced studies on integrating evaluation into strategic planning and governance in R&D (Milzow *et al*., 2019). At the same time, discussions on evaluation of S&T use in higher education intensified, especially around the UK's REF, which became a recurring reference. Studies by Bornmann (2013) Grant and Hinrichs (2015), Morgan Jones *et al*. (2017), and Parks *et al*. (2019) emphasized how universities used REF feedback to improve research impacts and strategies.

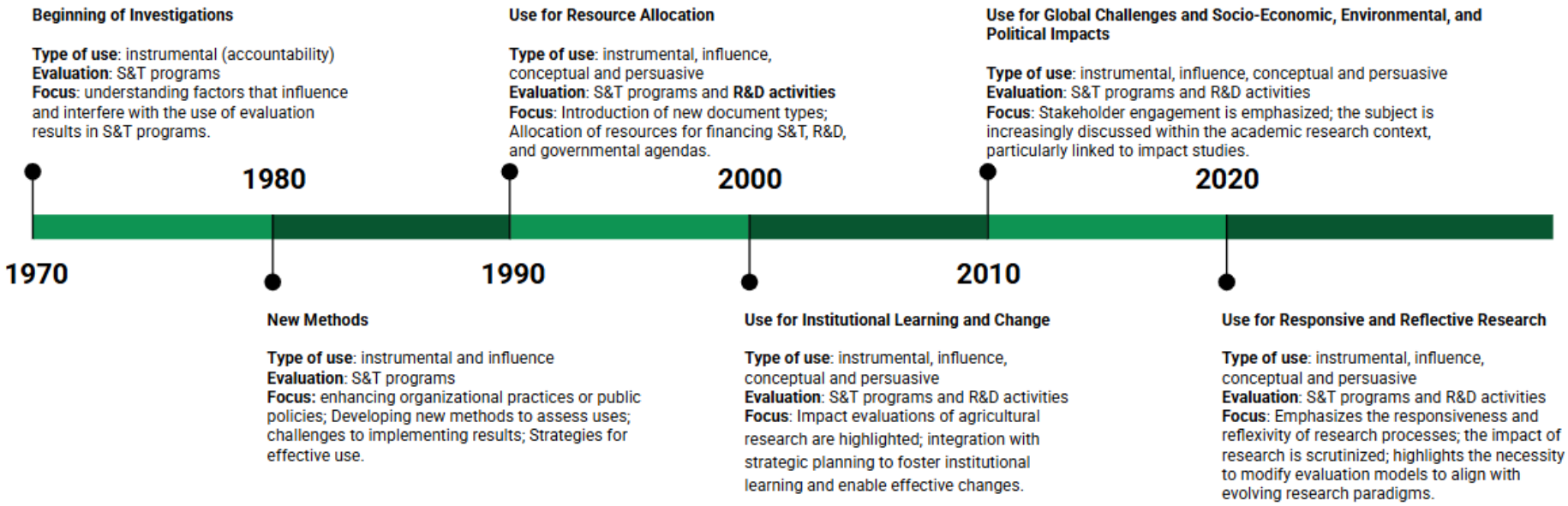
This decade also saw expanded dialogue on broader evaluation goals, promoting stakeholder participation, aligning evaluations with the "Europe 2020" strategy (Bayhan *et al*., 2011), reformulating R&D policies (Teirlinck *et al*., 2013), supporting rural development (Díez *et al*., 2016), informing public policy (Wollmann, 2016), and guiding agricultural research (Joly *et al*., 2016). Lee (2010) offered a distinct contribution, emphasizing the need for systematic management of evaluation information in S&T.

Between 1990 and 2010, persuasive/symbolic use emerged in 18 studies, 83% of which focused on S&T policies and programs. This form of use aims to validate or promote specific agendas, influencing stakeholders' opinions through strategic communication of results. It was especially noted in Social Sciences, Engineering, Agricultural, and Medical Sciences. Identified users included policymakers, program managers, and funding agencies.

From 2020 onward, studies have sustained the trend of examining evaluation use in academic research, now more closely tied to social, economic, and environmental impact. Stephens-Chu (2022) underscores universities' increasing reliance on evaluation results and the role of communication in facilitating their use. Brunet and Müller (2022) discuss how evaluation devices operate in competitive environments like the European Research Council, emphasizing the need for inclusive and adaptive evaluation practices. Eckhard and Jankauskas (2020) warn about political pressures that may compromise the integrity of evaluations in international organizations, while Kjesrud (2021) criticizes summative evaluation models for stifling innovation in teaching, calling for learning-oriented evaluation agendas.

Overall, the evolution from the 1970s to 2020 reveals a transition in the focus of studies, from resource allocation and accountability toward learning, reflexivity, strategy, and societal impact more recently (Figure 2). The four types of use described in Table 2, instrumental, conceptual, influence/process, and persuasive/symbolic, appear in different combinations across periods, knowledge fields, and user groups as shown in Table 7. This reflects the growing complexity and importance of evaluation as both a governance tool and a learning mechanism in S&T and R&D contexts, including users.

Figure 2 - Evolution of the discussion on the use of evaluation results in S&T and R&D



Source: The authors (2025).

Table 7 - Types of uses by knowledge fields and period

Evaluation	Decade	Field Oecd (2007)	Users	Instrumental Use	Conceptual Use	Persuasive Use	Influence Use
S&T Programs	1980	5. Social Sciences	Policymakers	2			
	1990	2. Engineering and Technology	Funding Agencies	2			2
			Programme Mannagers	2			
		5. Social Sciences	Funding Agencies	3			2
			Policymakers	3			
			Programme Mannagers	3	2		
	2000	2. Engineering and Technology	Industry	2	2		
			Policymakers	5	6	4	3
			Programme Mannagers	6	7	5	4
		3. Medical and Health Sciences	Policymakers	2	2	2	2
			Programme Mannagers	3	3	3	3
		4. Agricultural Sciences	Policymakers	3	3	3	3
			Programme Mannagers	4	4	4	4
		5. Social Sciences	Funding Agencies	2		2	2
			Policymakers	6	3	3	3
			Programme Mannagers	7	4	5	4
			Universities	2		2	
	2010	2. Engineering and Technology	Funding Agencies	3	2	2	2
			Industry	2	2		
			Policymakers	5	5	3	3
			Programme Mannagers	6	5	3	3
		3. Medical and Health Sciences	Health Institutes	2			
			Policymakers	4	2	2	
			Programme Mannagers	4	2	2	
		4. Agricultural Sciences	Policymakers	3	2	2	
			Programme Mannagers	3	2	2	
		5. Social Sciences	Academic Managers	4	4		
			Funding Agencies	2	2	2	2
			Health Institutes	2			
			Policymakers	9	7	5	4
			Programme Mannagers	10	7	4	4
			Universities	5	4		
R&D	1990	4. Agricultural Sciences	R&D Managers		2		
			R&D Organizations		2		
	2000	4. Agricultural Sciences	CGIAR	2	2		2
			R&D Managers	3	3		2
			R&D Organizations	3	3		2
	2010	2. Engineering and Techn..	Funding Agencies	2			
		3. Medical and Health Scie..	Funding Agencies	2			
		5. Social Sciences	Funding Agencies	2			

Source: The authors (2025).

Approaches and methods applied to investigate the use of evaluation results

In the majority of the analyzed sample, the topic is approached theoretically across all knowledge fields. In this regard, 33 studies (51%) adopted this approach, conducting literature reviews, documentary analysis, reflections, discussions, and propositions, respectively. This approach has not experienced many changes in terms of new methods over time, as we can observe in Table 8. However, it is noted that reflections and propositions have become more frequent from the second half of the 1990s, in both types of evaluation and particularly in the areas of Agricultural Sciences and Social Sciences.

It is worth highlighting the work of Patton and Horton (2009), who proposed a participatory model regarding UFE theory applied to agricultural R&D. Similarly, Lee (2010) proposes, through literature review and reflection, a conceptual model for the systematization of information management from evaluation results. Among some aspects, the model anticipates the identification of information from existing evaluation systems and the identification of information needs of potential users.

In this type of approach, techniques were applied to investigate the use of results from program and research evaluations for resource allocation, conceptual and theoretical understanding of use for social impact, accountability, barriers to use, learning, program planning, strategic management of R&D, organizational performance, influencing opinions and attitudes of stakeholders, as well as investigations into barriers to use.

Table 8 - Types of approaches and techniques employed

Approach	Evaluation	Decade	Techniques	Fields
Theoretical	S&T Programs	1970	literature review	5. Social Sciences
		1980	case studies; documental analysis	2. Engineering and Technology; 5. Social Sciences; 6. Humanities
		1990	discussion; documental analysis; reflexion	
		2000	discussion; documental analysis; literature review; reflexion	All six Areas
		2010	conceptual model; discussion; literature review; reflexion	
		2020	literature review	
	R&D	1990	discussion; literature review	All six Areas
		2000	discussion; literature review; reflection	4. Agricultural Sciences; 5. Social Sciences
Empirical	S&T Programs	1980	literature review; case studies; documental analysis; statistical analysis	1. Natural Sciences; 2. Engineering and Technology; 3. Medical and Health Sciences; 5. Social Sciences
		1990	case studies; documental analysis; interviews	All six Areas
		2000	case studies	2. Engineering and Technology; 5. Social Sciences
		2010	case studies; focal groups & observation; interviews; surveys	All six Areas
		2020	case studies	5. Social Sciences
	R&D	1990	interviews	1. Natural Sciences; 2. Engineering and Technology; 3. Medical and Health Sciences; 5. Social Sciences
		2000	surveys	4. Agricultural Sciences
		2010	case studies; text mining	All six Areas
Mixed	S&T Programs	1980	interviews; literature review	5. Social Sciences
		1990	interviews; statistical analysis	3. Medical and Health Sciences; 5. Social Sciences
	S&T Programs	2000	case studies; comparative analysis; component analysis; documental analysis; focal groups & observation; interviews; literature review; statistical analysis	All six Areas
	S&T Programs	2010	case studies; focal groups & observation; literature review; statistical analysis; surveys	2. Engineering and Technology; 4. Agricultural Sciences; 5. Social Sciences; 6. Humanities
	R&D	2010	case studies; documental analysis; literature review	All six Areas, but Agricultural Sciences

Source: The authors (2025).

Seventeen studies (27%) are situated within the empirical approach, where the predominant methods are case studies and interviews, reflecting a need to deeply understand the contexts and impacts of S&T programs and research. Surveys complement these data collection approaches, in addition to techniques of direct observation by the researcher. Empirical studies began from the 1980s to 2020 with progressive growth and are present in more recent publications across all areas for investigations on the use in S&T programs. Interviews were employed in studies from the 1990s and 2010s, while surveys were used in the 2000s and 2010s. Specifically, in investigations on the use in R&D evaluations, new techniques were identified, such as text mining and specific quantitative analyses (like Poisson regression), which are associated with the complexity and specificity of R&D data, such as those that aimed to ascertain the use of the results of academic research impact evaluations for resource allocation from the REF (Grant; Hinrichs, 2015).

Twelve studies incorporated approaches that blended theoretical and empirical techniques. These began in the 1980s within the Social Sciences domain, particularly in studies related to the use within S&T program contexts. By the 1990s, there was a gradual increase in the use of this approach within the fields of Engineering and Technology, followed by Natural and Social Sciences. In R&D, mixed-method studies emerged in the 2010s, utilizing case studies, literature reviews, and document analysis, particularly in Natural Sciences, Engineering and Technology, and Social Sciences. The volume of these studies significantly increased from the 2000s onward. These investigations often employed literature review combined with case studies techniques to explore barriers to use, resource allocation, program planning, and decision-making processes.

Similarity among the studies

In verifying the similarity among the studies, through topic modeling, four groups were created and helped to organize the theme regarding S&T and R&D. These groups represent, in an organized manner, the studies that share the same kind of document type and semantic similarities related to the theme and subject discussed (measured by the presence of common words (tokens)), and even methodological choices. It is important to note that the same work can belong to more than one group, as seen in Table 9.

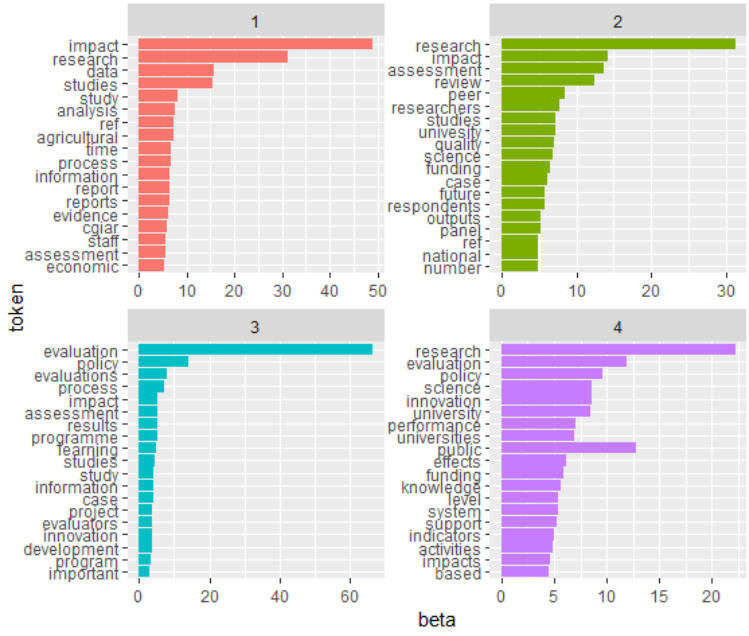
Table 9 - Similarity between studies

Group	Works	Amount
1	Most2010.pdf; nelson_2005.pdf; parks_2019.pdf; patton_horton2009.pdf; pollard1983.pdf; raitzer2008.pdf; ramaligam_2011.pdf; roseland2011.pdf; thompson1983.pdf; horton1998.pdf; horton2003.pdf	11
2	Bayhan_2011.pdf; bornmann2013.pdf; flemming1991.pdf; fteval_2003.pdf; grant_2015.pdf; hall1981.pdf; hansson2006.pdf; johnson2009.pdf; manville_2015.pdf; morgan_2017.pdf; roseland2011.pdf; sans_menendez_1997.pdf	12
3	sans_menendez_1997.pdf; barker2007.pdf; Bayhan_2011.pdf; boyer1991.pdf; brofoss1998.pdf; brunet2022.pdf; Carlsson-1999-Are.pdf; chelimsky_2001.pdf; Diez2016.pdf; flemming1991.pdf; fteval_2003.pdf; hall1981.pdf; hansson2006.pdf; hemlin1996.pdf; joly2016.pdf; jordan2008.pdf; Kamenetzky_2016.pdf; kjesrud2021.pdf; Kuhlmann_2001.pdf; kuhlmann1998.pdf; luukkonen1995.pdf; macdonald2006.pdf; mackay1992conv.pdf; mackay2003.pdf; manville_2015.pdf; meyer-krahmer1989.pdf; milzow2018.pdf	27
4	amanatidou_2014.pdf; baker2013.pdf; balthasar_2000.pdf; barker2007.pdf; Bayhan_2011.pdf; bennet_cgiar_2001.pdf; bornmann2013.pdf; boyer1991.pdf; brofoss1998.pdf; brunet2022.pdf; Carlsson-1999-Are.pdf; chelimsky_2001.pdf; chelimsky1977.pdf; cozzens1997.pdf; Diez2016.pdf; grant_2015.pdf; Gök_2010.pdf; gok_2012.pdf; joly2016.pdf; jordan2008.pdf; Kamenetzky_2016.pdf; kjesrud2021.pdf; kuhlmann1998.pdf; luukkonen1995.pdf; macdonald2006.pdf; mackay1992conv.pdf; mackay2003.pdf; manville_2015.pdf; meyer-krahmer1989.pdf; milzow2018.pdf; sans_menendez_1997.pdf; saunders_2005.pdf; shapira_2001.pdf; siegel1985.pdf; smith_2007.pdf; stephens-chu2022.pdf; Technopolis_2010.pdf; teirlinck2012.pdf; ter_bogt_2009.pdf; thakur2007.pdf; thompson1983.pdf; vakola2000.pdf; Wollmann_2016.pdf	43

Source: The authors (2025).

According to Figure 3, Group 1 comprises studies that analyzed the use of evaluation results more focused on R&D evaluation. Tokens such as "impact," "research," combined with others like "agricultural," "cgiar," and "ref," confirm this evidence. Consequently, this is the group with the fewest publications. It mainly features reports that address the theme through case studies. The agricultural area, with CGIAR being the most represented organization, and the social sciences area, evidenced by the presence of REF, are the most prominent in this group.

Figure 3 - Topic groups created from the sample



Source: The authors (2025).

In Group 2, the studies also focus on the use of impact evaluation results in R&D, with tokens like "research" and "impact" recurring. This group, although the second smallest, includes more studies associated with program evaluation and national systems for evaluating academic research. This is evidenced by tokens such as "university," "public," "funding," "quality," and "national," indicating a focus on resource allocation related to funding agencies and public research financing, with the REF often cited as a common reference.

Group 3 discusses how evaluation results support program adjustments and the development of new S&T policies and programs and R&D projects and activities. These studies explore how evaluation outcomes guide strategic decisions, enhance organizational learning, and drive changes aimed at meeting global commitments and challenges. Tokens like "influence," "change," and "decisions" highlight this focus on responsible, innovation-driven research tailored to contemporary and future needs.

Finally, Group 4 contains the largest number of publications, primarily focusing on the use of results in S&T programs to optimize resource allocation and maximize impacts. Tokens such as "funding," "projects," and "indicators" reflect concerns about the efficiency and effectiveness of S&T investments. At the same time, tokens like "learning" and "innovation" indicate an orientation towards using results for organizational learning and fostering a culture of innovation.

Final Considerations

This study systematically explored the literature on the use of evaluation results in S&T and R&D and examined its evolution over time. The topic has been more extensively studied in S&T policies and programs since the 1970s, showing consistent growth. In R&D, it emerged in the 1990s, expanded in the 2000s, declined in the 2010s, and shows renewed interest in the 2020s.

These differences reflect structural characteristics of the two contexts. S&T initiatives have broader scope and visibility, generating stronger accountability pressures, whereas R&D projects are more internally oriented and only recently have faced demands to demonstrate societal value (Spaapen, 2015; Reed *et al*., 2021; 2022). This shift aligns with expectations that research produces social, environmental, economic, and political impacts, as seen in assessments such as the UK REF (Morgan *et al*., 2017; Parks *et al*., 2019), and is reinforced by RRI and RRA approaches (Felt *et*

al., 2018). Accordingly, evaluation use has expanded beyond resource allocation to include learning and strategic alignment.

Over the decades, efforts to investigate how evaluation results are used have grown, especially since the 1980s in S&T and the 1990s in R&D. These studies reflect an increasing interest in leveraging evaluation for institutional learning and addressing global challenges. Laredo and Mustar (2001) noted that science and research in Western countries underwent significant funding cuts in the 1990s, which may explain the intensified focus on efficiency and impact. The fields that gained prominence: Social Sciences, Agriculture, Engineering, and Medicine, are not only fundamental but closely linked to global challenges, such as food security and public health.

Agriculture, for example, plays a central role in climate and food system discussions, while Medicine was particularly visible during the COVID-19 pandemic. In R&D, initial discussions on use of evaluation results often centered on impact studies in Agriculture, especially those from CGIAR, as demonstrated by Horton (1998), Horton and Mackay (2003), and Mackay and Horton (2003), who emphasized evaluations as tools for strategy and learning.

Conversely, Natural Sciences and Humanities were not strongly represented in the sample. When they did appear, it was usually in broader discussions encompassing all fields, such as those by Chelimsky (2001), Grant and Hinrichs (2015), Ter Bogt and Scapens (2009), and Shapira and Kuhlmann (2003). Their limited presence may reflect the more abstract nature and longer-term impacts of work in these areas, as well as weaker institutional mechanisms for evaluation use.

An imbalance was also observed between the Global North and South. Most studies originated in countries with long-standing traditions in evaluation, such as those in Europe and the United States (Cruz Castro; Sanz Menéndez, 2005; 2006; Sanz Menéndez, 1997). In the Global South, as Stockmann *et al*. (2022) and Maldonado Trujillo (2022) argue, there is still a need to establish and institutionalize evaluation practices, particularly through ECB. As these capacities grow, it becomes essential to understand how evaluation results are being used in policy and organizational contexts in these regions.

Another relevant gap is the limited systematization of how evaluation results are used. Lee (2010) emphasizes the need to treat evaluations as structured information sources, proposing a model to manage evaluation data and stimulate its application.

This includes linking S&T policy evaluations with R&D assessments to create more integrated and responsible evaluation ecosystems.

In this context, the "Branch of Use" (Alkin; Christie, 2004; 2023) offers conceptual and operational foundations for understanding and improving the use of evaluation results. It brings together theoretical frameworks and practical tools capable of addressing the specific demands of S&T and R&D contexts. While not the primary focus of this work, many studies cited key figures in this tradition, such as Weiss (1998), Patton (2008; 2015), and Cousins and Leithwood (1986), indicating its strong influence.

The findings of this review demonstrate that while S&T evaluation results may serve broader institutional sets, those from R&D tend to be more internal, shaped by organizational cultures and needs. Morgan *et al*. (2017) and Parks *et al*. (2019), for instance, highlight how REF evaluation results have supported decision-making in English universities. Similarly, Milzow *et al*. (2019) point out the relevance of organizational structures and leadership in shaping how evaluations are used in R&D environments.

Methodologically, the combination of the PRISMA protocol with content analysis and topic modeling enabled a structured mapping of the literature. By distinguishing S&T from R&D and grouping thematic similarities computationally, the study offers a replicable framework for investigating evaluation use. The four document groups identified organize trends, methods, and content patterns, supporting more focused inquiries. For policymakers and R&D managers, the findings reinforce evaluation as a strategic resource integrated into planning and decision-making, aligned with RRI and RRA principles. Thus, evaluation should function as a dynamic instrument that informs and guides action, rather than a merely symbolic exercise.

In terms of limitations, the analysis was restricted to documents indexed in Web of Science, Scopus, and Google Scholar, which may underrepresent publications from local or institutional outlets, particularly in the Global South. Second, the classification of knowledge fields and types of use relied on available descriptions in the articles, which may introduce interpretative bias. Third, the use of automated text analysis techniques, while enabling large-scale mapping, may simplify conceptual nuances present in qualitative discussions. Finally, the temporal distribution of publications affects the identification of trends, especially in more recent years. These limitations open opportunities for future research. Further studies could incorporate grey literature

and institutional reports, conduct comparative analyses across regions, and investigate empirically how evaluation results are used within organizations. In particular, qualitative and mixed-method approaches may deepen the understanding of organizational processes that mediate the transition from evaluation production to evaluation use.

References

ALKIN, M. C. *et al*. **Using evaluations**: does evaluation make a difference?. [S. l.]: Guilford Publications, 1979.

ALKIN, M. C.; CHRISTIE, C. A. An evaluation theory tree. **Evaluation roots**: tracing theorists' views and influences, [S. l.], v. 2, n. 19, p. 12–65, 2004.

ALKIN, M. C.; CHRISTIE, C. A. **Evaluation Roots**: theory influencing practice. [S. l.]: Guilford Publications, 2023.

ALKIN, M. C.; KING, J. A. The historical development of evaluation use. **American Journal of Evaluation**, [S. l.], v. 37, n. 4, p. 568–579, dez. 2016.

AMANATIDOU, E. *et al*. Using evaluation research as a means for policy analysis in a 'new'mission-oriented policy context. **Minerva**, [S. l.], v. 52, p. 419–438, 2014.

BAYHAN, D. *et al*. **Optimizing the research and innovation policy mix**: the practice and challenges of impact assessment in Europe. [S. l.: S. n.], 2011. Disponível em: https://kuleuven.limo.libis.be/discovery/fulldisplay/lirias1679390/32KUL_KUL:Lirias. Acesso em: 10 jul. 2024.

BORNMANN, L. What is societal impact of research and how can it be assessed?: a literature survey. **Journal of the American Society for information science and technology**, [S. l.], v. 64, n. 2, p. 217–233, 2013.

BOURGEOIS, I.; COUSINS, J. B. Informing evaluation capacity building through profiling organizational capacity for evaluation: an empirical examination of four Canadian Federal Government Organizations. **Canadian Journal of Program Evaluation**, [S. l.], v. 23, n. 3, p. 127–146, jan. 2009.

BOZEMAN, B.; SAREWITZ, D. Public Value mapping and science policy evaluation. **Minerva**, [S. l.], v. 49, n. 1, p. 1–23, mar. 2011.

BRUNET, L.; MÜLLER, R. Making the cut: how panel reviewers use evaluation devices to select applications at the European Research Council. **Research Evaluation**, [S. l.], v. 31, n. 4, p. 486–497, 2022.

CHELIMSKY, E. Improving the use of evaluation: an agency perspective. **The Prison Journal**, [S. l.], v. 57, n. 1, p. 13–18, abr. 1977.

CHELIMSKY, E. What evaluation could do to support foundations: a framework with nine component parts. **American Journal of Evaluation**, [S. l.], v. 22, n. 1, p. 13–28, 2001.

COARA. **Agreement on reforming research assessment**: now open for signature. [S. l.: S.n.], 2022. Disponível em: https://research-and-innovation.ec.europa.eu/news/all-research-and-innovation-news/agreement-reforming-research-assessment-now-open-signature-2022-10-03_en. Acesso em: 5 jun. 2023.

COUSINS, J. B. *et al*. Organizational capacity to do and use evaluation: results of a pan-canadian survey of evaluators. **Canadian Journal of Program Evaluation**, [S. l.], v. 23, n. 3, p. 1–35, jan. 2009.

COUSINS, J. B.; LEITHWOOD, K. A. Current empirical research on evaluation utilization. **Review of educational research**, [S. l.], v. 56, n. 3, p. 331–364, 1986.

COZZENS, S. E. Frontiers of evaluation: United States. **Revista Brasileira de Inovação**, [S. l.], v. 11, p. 47–66, 30 jul. 2012.

COZZENS, S. E.; MELKERS, J. E. Use and usefulness of performance measurement in state science and technology programs. **Policy Studies Journal**, [S. l.], v. 25, n. 3, p. 425–435, set. 1997.

CRUZ CASTRO, L.; SANZ MENÉNDEZ, L. **Research Evaluation in transition**: individual versus organizational assessment in Spain. Madrid: Uout, 2006.

CRUZ-CASTRO, L.; SANZ-MENÉNDEZ, L. Politics and institutions: European parliamentary technology assessment. **Technological Forecasting and Social Change**, [S. l.], v. 72, n. 4, p. 429–448, maio, 2005.

DÍEZ, M. A.; IZQUIERDO, B.; MALAGÓN, E. Increasing the use of evaluation through participation: the experience of a rural sustainable development plan evaluation. **Environmental Policy and Governance**, [S. l.], v. 26, n. 5, p. 366–376, set. 2016.

DONNELLY, C. *et al*. Supporting knowledge translation through evaluation: evaluator as knowledge broker. **Canadian Journal of Program Evaluation**, [S. l.], v. 29, n. 1, p. 36–61, mar. 2014.

ECKHARD, S.; JANKAUSKAS, V. Explaining the political use of evaluation in international organizations. **Policy Sciences**, [S. l.], v. 53, n. 4, p. 667–695, dez. 2020.

FELT, U. **Responsible research and innovation**. New York: Routledge, 2018.

FORAY, D.; MOWERY, D. C.; NELSON, R. R. Public R&D and social challenges: What lessons from mission R&D programs?. **Research policy**, [S. l.], v. 41, n. 10, p. 1697–1702, 2012.

GALVÃO, T. F.; PANSANI, T. de S. A.; HARRAD, D. Principais itens para relatar revisões sistemáticas e meta-análises: a recomendação PRISMA. **Epidemiologia e Serviços de Saúde**, [S. l.], v. 24, p. 335–342, jun. 2015.

GRANT, J.; HINRICHS, S. **The nature, scale and beneficiaries of research impact**: an initial analysis of the Research Excellence Framework (REF) 2014 impact case studies. [S. l.: S. e.], 2015.

HALL, G. E. **Viewing Evaluation Utilization as an Innovation**. [S. l.: S. e.], 1981.

HARZING, A.W. Publish or Perish. **Harzing.com**. Research in International Management. Lodon, 26 fev. 2025. Disponível em:

https://harzing.com/resources/publish-or-perish. Acesso em: 9 jul. 2024.

HARZING, A. W.; ALAKANGAS, S. Google Scholar, Scopus and the web of science: a longitudinal and cross-disciplinary comparison. **Scientometrics**, [S. l.], v. 106, n. 2, p. 787–804, fev. 2016.

HEMLIN, S. Research on research evaluation. **Social Epistemology**, [S. l.], v. 10, n. 2, p. 209–250, abr. 1996.

HORTON, D. Disciplinary roots and branches of evaluation: Some lessons from agricultural research. **Knowledge and Policy**, [S. l.], v. 10, n. 4, p. 31–66, dez. 1998.

HORTON, D. Book review: utilization-focused evaluation. **American Journal of Evaluation**, [S. l.], v. 20, n. 2, p. 399–401, 1999.

HORTON, D.; MACKAY, R. Using evaluation to enhance institutional learning and change: recent experiences with agricultural research and development. **Agricultural Systems**, [S. l.], v. 78, n. 2, p. 127–142, 2003.

JOHNSON, K. *et al*. Research on Evaluation Use: A Review of the Empirical Literature From 1986 to 2005. **American Journal of Evaluation**, [S. l.], v. 30, n. 3, p. 377–410, 1 set. 2009.

JOLY, P. B. *et al.* **Agricultural research impact assessment**: issues, methods and challenges. [S. l.: S. e.], 2016.

KJESRUD, R. D. **Using Assessment to Prompt Innovation**. [S. l.: S. e.], 2021.

KOUSHA, K.; THELWALL, M. Google Scholar citations and Google Web/URL citations: a multi-discipline exploratory analysis. **Journal of the American Society for Information Science and Technology**, [S. l.], v. 58, n. 7, p. 1055–1065, 2007.

KUHLMANN, S. Moderation of Policy-Making?: science and technology policy evaluation beyond impact measurement: the case of Germany. **Evaluation**, [S. l.], v. 4, n. 2, p. 130–148, abr. 1998.

LABIN, S. N. *et al*. A research synthesis of the evaluation capacity building literature. **American Journal of Evaluation**, [S. l.], v. 33, n. 3, p. 307–338, 2012.

LARÉDO, P.; MUSTAR, P. **Research and innovation policies in the new global economy**: an international comparative analysis. [S. l.]: Edward Elgar Publishing, 2001.

LEE, J. **How to manage and utilize the information from R&D evaluation?**. [S. l.]: IEEE, 2010.

LEVITON, L. C.; HUGHES, E. F. X. Research on the utilization of evaluations: A review and synthesis. **Evaluation review**, [S. l.], v. 5, n. 4, p. 525–548, 1981.

LUUKKONEN, T. The impacts of research field evaluations on research practice.

Research Policy, [S. l.], v. 24, n. 3, p. 349–365, 1995.

MACKAY, K. The use of evaluation in the budget process. **Australian Journal of Public Administration**, [S. l.], v. 51, n. 4, p. 436–439, 1992.

MACKAY, R.; HORTON, D. Expanding the use of impact assessment and evaluation in agricultural research and development. **Agricultural Systems**, [S. l.], v. 78, p. 143–165, 1 nov. 2003.

MALDONADO TRUJILLO, C. Centers for Learning on Evaluation and Results. Latin America and the Caribbean. *In*: **The institutionalisation of evaluation in the Americas**. [S. l.]: Springer, 2022. p. 381–401.

MILZOW, K. *et al*. Understanding the use and usability of research evaluation studies. **Research Evaluation**, [S. l.], v. 28, n. 1, p. 94–107, 2019.

MOHER, D. *et al*. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. **Annals of internal medicine**, [S. l.], v. 151, n. 4, p. 264–269, 2009.

MONGEON, P.; PAUL-HUS, A. The journal coverage of Web of Science and Scopus: a comparative analysis. **Scientometrics**, [S. l.], v. 106, p. 213–228, 2016.

MORGAN JONES, M.; GRANT, J.; RAND, E. **Making the grade:** methodologies for assessing and evidencing research impact. [S. l.]: Dean, Wykes and Stevens editors, 2013.

MORGAN JONES, M.; MANVILLE, C.; CHATAWAY, J. Learning from the UK's research impact assessment exercise: a case study of a retrospective impact assessment exercise and questions for the future. **The Journal of Technology Transfer**, [S. l.], p. 1–25, 2017.

NATIONAL SCIENCE FOUNDATION. **EvaluATE**: ATE Evaluation Resource Hub. Michigan: Kalamazzo, 2024. Disponível em: https://evalu-ate.org/. Acesso em: 26 mar. 2024.

NORTHAN, J. **The Research Excellence Framework (REF)**: [S. l.]: The BU REF Project, 2010.

ORGANIZATION FOR ECONOMIC COOPERATION AND DEVELOPMENT. Fields of science and technology classifications. **EduTech Wiki**, [S. l.], 2013. Disponível em: https://edutechwiki.unige.ch/en/Fields_of_science_and_technology_classifications. Acesso em: 30 nov. 2023.

OLEJNICZAK, K. The Game of Knowledge Brokering: a new method for increasing evaluation use. **American Journal of Evaluation**, [S. l.], v. 38, n. 4, p. 554–576, dez. 2017.

OLEJNICZAK, K.; RAIMONDO, E.; KUPIEC, T. Evaluation units as knowledge brokers: Testing and calibrating an innovative framework. **Evaluation**, [S. l.], v. 22, n. 2, p. 168–189, abr. 2016.

PAGE, M. J. *et al*. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. **International journal of surgery**, [S. l.], v. 88, p. 105906, 2021a.

PAGE, M. J. *et al*. Updating guidance for reporting systematic reviews: development of the PRISMA 2020 statement. **Journal of clinical epidemiology**, [S. l.], v. 134, p. 103–112, 2021b.

PARKS, S. *et al*. **The changing research landscape and reflections on national research assessment in the future**. [S. l.]: RAND, 2019.

PATTON, M. Q. **Utilization-focused evaluation**. [S. l.]: Sage publications, 2008.

PATTON, M. Q. The sociological roots of utilization-focused evaluation. **The American Sociologist**, [S. l.], v. 46, n. 4, p. 457–462, 2015.

PATTON, M. Q.; HORTON, D. **Utilization-focused evaluation for agricultural innovation**. [S. l.], ILAC Brief, 2009.

PINTO, D. M.; BIN, A.; RODRIGUES, G. S. Trends and challenges in the evaluation of agricultural research impact: a systematic literature review. **Cadernos de Ciência & Tecnologia**, Brasília, v. 42, e27611, 2025. DOI: https://doi. org/10.35977/0104-1096.cct2025.v42.27611. Disponível em: https://apct.sede.embrapa.br/cct/article/view/27611. Acesso em: 8 jul. 2024.

PRESKILL, H.; BOYLE, S. A multidisciplinary model of evaluation capacity building. **American journal of evaluation**, [S. l.], v. 29, n. 4, p. 443–459, 2008.

PRESKILL, H.; BOYLE, S. Insights into evaluation capacity building: motivations, strategies, outcomes, and lessons learned. **Canadian Journal of Program Evaluation**, [S. l.], v. 23, n. 3, p. 147–174, jan. 2009.

RAITZER, D. A.; KELLEY, T. G. Assessing the contribution of impact assessment to donor decisions for international agricultural research. **Research Evaluation**, [S. l.], v. 17, n. 3, p. 187–199, 2008.

REED, M. S. *et al.* Evaluating impact from research: a methodological framework. **Research Policy**, [S. l.], v. 50, n. 4, 2021.

REED, M. S. *et al*. How can impact strategies be developed that better support universities to address twenty-first-century challenges? **Research for All**, [S. l.], v. 6, n. 1, 2022.

RESEARCH EXCELLENCE FRAMEWORK. Securing a world-class, dynamic and responsive research base across the full academic spectrum within UK higher education. **REF 2029**. United Kingdom, 2024. Disponível em: https://www.ref.ac.uk/. Acesso em: 8 jul. 2024.

SANZ MENÉNDEZ, L. **Science and technology policy evaluation in the context of advanced S&T policy planning**. [S. l.: S. e.], 1997.

SCHÖNBRODT, F. *et al*. **Responsible Research Assessment I**: implementing DORA for hiring and promotion in psychology. [S. l.: S. e.], 2022.

SHAPIRA, P.; KUHLMANN, S. Learning from science and technology policy evaluation. *In:* **Learning from Science and Technology Policy Evaluation**. [S. l.]: Edward Elgar Publishing, 2003. p. 1–17.

SIEGEL, K.; TUCKEL, P. The utilization of evaluation research: a case analysis. **Evaluation Review**, [S. l.], v. 9, n. 3, p. 307–328, jun. 1985.

SILGE, J.; ROBINSON, D. **Text mining with R**: a tidy approach. First edition ed. Beijing; Boston: O'Reilly, 2017.

SPAAPEN, J. B. A new evaluation culture is inevitable. **Organic Farming**, [S. l.], v. 1, n. 1, p. 36–37, 2015.

STEPHENS-CHU, M. **Using assessment results at a large research university**: methods of analysis and exemplary practices. [S. l.: S. e.], 2022.

STILGOE, J.; GUSTON, D. **Responsible research and innovation**. [S. l.]: MIT Press, 2016.

STOCKMANN, R.; MEYER, W.; SZENTMARJAY, L. **The institutionalisation of evaluation in the Americas**. [S. l.]: Springer Nature, 2022.

STUFFLEBEAM, D. L. **Evaluation as enlightenment for decision-making**. 1968.

STUFFLEBEAM, D. L. The CIPP model for Program Evaluation. *In*: MADAUS, G. F.; SCRIVEN, M. S.; STUFFLEBEAM, D. L. (Ed.). **Evaluation Models**. Dordrecht: Springer Netherlands, 1983. p. 117–141.

STUFFLEBEAM, D. L. The CIPP Model for Evaluation. In: STUFFLEBEAM, D.; MADAUS, G. F.; KELLAGHAN, T. (org.). **Evaluation Models**: evaluation in education and human services. Dordrecht: Kluwer Academic Publishers, 2002.

TEIRLINCK, P. *et al*. Closing the policy cycle: increasing the utilization of evaluation findings in research, technological development and innovation policy design. **Science and Public Policy**, [S. l.], v. 40, n. 3, p. 366–377, 2013.

TER BOGT, H. J.; SCAPENS, R. W. Performance measurement in universities: a comparative study of two A&F groups in the Netherlands and the UK. **Available at SSRN**. [S. l.], v. 1, 2009.

THOMPSON, B.; LEVITOV, J. E. Quantitative analysis of previous research on use of evaluation. **Psychological Reports**, [S. l.], v. 53, n. 1, p. 215–221, ago. 1983.

VAN DER MOST, F. Use and non-use of research evaluation: a literature review. **CIRCLE Paper**, [S. l.], v. 1, n. 16, 2010.

VON SCHOMBERG, R. Why responsible innovation. **International handbook on**

responsible innovation: a global resource, [S. l.], n. 1, p. 12–32, 2019.

WEISS, C. H. Utilization of evaluation: toward comparative study. **Evaluating action programs: Readings in social action and education**, [S. l.], n. 1, p. 318–326, 1972.

WEISS, C. H. Have we learned anything new about the use of evaluation?. **The American Journal of Evaluation**, [S. l.], v. 19, n. 1, p. 21–33, 1998.

WHOLEY, J. S. Evaluability assessment. **Handbook of practical program evaluation**, [S. l.], v. 2, p. 33–62, 2004.

WOLLMANN, H. Utilization of evaluation results in Policy-Making Aad Administration: a challenge to Political Science Research. **Hrvatska i komparativna javna uprava: časopis za teoriju i praksu javne uprave**, [S. l.], v. 16, n. 3, p. 433–458, 2016.