

## EVALUATION: DEFINITION, PURPOSE AND USES

Ángel Mundo López<sup>1</sup>

<sup>1</sup> Profesor-investigador de la Universidad Autónoma Metropolitana Unidad Xochimilco, adscrito al Departamento de Política y Cultura.

\*Author for correspondence: [amundo@correo.xoc.uam.mx](mailto:amundo@correo.xoc.uam.mx)

---

### ABSTRACT

The document addressed the evaluation of public policies, their relevance and how to optimize their implementation to improve social programs. A methodology based on theories from authors such as Weiss and Cohen were used; qualitative and quantitative evaluation models were analyzed; it also included experimental and quasi-experimental techniques. Evaluation was defined as an activity aimed at making judgments based on specific criteria, with the objective of identifying achievements and areas of improvement in social programs. Additionally, the “implementation gap” is discussed and we analyze the impact of the evaluation on accountability and budget control. Impact evaluations favor experimental methods, but it was observed how qualitative methods, in combination, help to better understand the reasons behind the results. The study concludes that, although evaluation should contribute to decision-making, its success depends on an appropriate political context and a flexible approach that considers the realities of each program. It is emphasized that evaluation should not be used as an end, but as a strategic tool that allows optimizing policies and practices, recognizing the importance of adapting its approaches to the concrete and social situation of each intervention. Based on this analysis, it is proposed that policy evaluation, when used appropriately and thoughtfully, is a fundamental tool to promote transparency and effectiveness in government interventions, contributing to the exercise of citizen rights and the improvement of social well-being.

**Keywords:** accountability, programmatic effectiveness, decision making.

---

### INTRODUCTION

Policy evaluation acquires, with the passage of time, increasing relevance; No one currently denies, at least in public (Thoenig, 2018), the advantages offered by evaluating policies; However, this public praise does not correlate with issues such as promotion and financing, not to mention utilization, which are issues that are still

**Citation:** Mundo-López  
A. 2025. Evaluation:  
Definition, purpose and  
uses.

**REMEVAL** 1(1): 46-57.  
[https://doi.org/10.63121/  
fw29fs20](https://doi.org/10.63121/fw29fs20)

**Received:**  
12 September, 2024

**Accepted:**  
18 October, 2024

**Published:**  
12 February, 2025

This work is licensed  
under a Creative  
Commons Attribution-  
Non- Commercial 4.0  
International license.



pending<sup>2</sup>. Despite the laudatory speech, ambiguities persist regarding its meaning, uses and exploitation, therefore, without the intention of clarifying an issue that has been in the foreground for decades, we will make our contribution to the debate in a synthetic way.

First, we will focus on the definition of the term evaluation, because, although it seems like a common sense issue, talking about the evaluation of policies or programs requires a greater specification; We will continue with the characteristics that an evaluation is expected to have, and for that we will rely on the proposal of Weiss (2018) that establishes five characteristics that must be met, trying to delve into its explanations by adapting them to the Mexican context. Finally, we will reach some brief conclusions.

## THEORETICAL FRAMEWORK

According to Weiss (2018), throughout our lives we carry out acts of evaluation: such as when we rate a service we receive, when we judge the usefulness of a product, or when we make a judgment about the seasoning of some food that we receive. we taste Something similar is commented by Meny & Thoenig (1992), who say that: “Evaluation is a daily gesture [...] In their own way, all citizens make judgments about what the government of a State or the city council does.”

The most common way in which as children we are linked to the topic of evaluation is with the application of questionnaires or tests in which the level of achievement achieved in school lessons is “judged”, that is, after setting a standard ( a set of expected learning), a series of lessons is carried out that are expected to be assimilated by the students, and to estimate whether the objective has been achieved, those instruments are used. However, these types of “evaluations” generally take on a punitive nature, in which the person who does not reach the minimum requirements is forced to retake the course until what is expected is “met”; and although the establishment of formative (and therefore non-punitive) evaluation mechanisms has been advocated for some years now, work still needs to be done on their design and implementation.

As can be seen, evaluation is a term that accompanies us from our first stages of life; However, in the case of the evaluation of policies and programs, particularly with a social objective, the definition of this activity takes on a slightly more technical nature. Cohen & Franco (1992) define evaluation as:

---

<sup>2</sup> The Mexican meta-evaluations also identify important problems in the use of the recommendations arising from the evaluation processes (Cardozo, 2012).

*“The attempt to compare a pattern of desirability (objective-image towards which the action is oriented) with reality (the potential extent to which it will be modified, or what really happened as a consequence of the activity deployed) and, on the other hand, the concern to effectively achieve the stated objectives [...] To evaluate is to establish the value of a thing; To do so, a procedure is required by which what is to be evaluated is compared with respect to a specific criterion or pattern.”*

Based on the above, we can propose a definition like the following: evaluation is the issuance of a judgment on some public intervention, which is issued after having applied different research techniques, both qualitative and quantitative, that allow analyzing the scope of the program objectives, to reach a set of conclusions and issue recommendations that may eventually help improve the effectiveness and operation of said intervention.

In this definition proposal we try to recover some of the points that Weiss (2018) postulates that the evaluation must comply with: 1) Carrying out a systematic assessment, which is issued, either, 2) on the operation or, 3) on the impacts. of the program based on, 4) the establishment of certain standards that serve to make the comparison between what is sought and what is desired to be achieved, finally, 5) identifies the purpose of carrying out the evaluation. We will develop each of these points below.

Issuance of a value judgment that is arrived at thanks to the application of a certain method. Although since the end of the last century, Patton reflected on the “war of paradigms” which, from his perspective, did not help the development of evaluation, and therefore, he advocated the establishment of combined approaches, the reality is that impact evaluations have a bias strongly aligned with quantitative methods, specifically through the application of experimental processes (despite the existing complications, such as, for example, the greater time required for the application of observation techniques as well as a greater amount of economic resources; ethical -deliberately leaving sectors that need it without attention; guaranteeing validity, both external and internal, etc.). Given this, Campbell & Stanley (1973) say:

*“There are many social situations in which the researcher can introduce something similar to the experimental design in his programming of procedures (e.g., the when and to whom of the measurement), although he lacks complete control over the programming of experimental stimuli ( the when and to whom of the exposure and the ability to randomize it), which allows an authentic experiment to be carried out [...] because there is a lack of total experimental control, it is essential that the researcher has in-depth knowledge of which are the specific variables that its particular design does not control”.*

Thanks to the flexibility of quasi-experimental methods, the rigor that must be taken to control the sources of invalidation, both internal (history, maturation, test administration, etc.) and external (reactive or interaction effect of the tests, effects of interaction of selection biases and the experimental variable, etc.)<sup>3</sup>.

The important thing, in any case, is to identify what has been achieved (or why the desired objectives have not been obtained) through the application of a methodology that, ideally, can combine both qualitative and quantitative research techniques. Following Deleau *et al.* (1986), the aim is to answer questions such as: To what degree has public action caused the objectives to be achieved? Which of its elements have contributed? Could another policy have been more effective? Are these results generalizable to policies of the same nature, but with target populations or applied in different regions? Has the policy had positive or negative effects? What can be said ex-post about the relevance of its objectives or what other objectives should replace the current ones?

Next, Weiss refers to the way in which the policy is being put into practice, that is, the operation (instrumentation or implementation). From the landmark study by Pressman & Wildavski (1998), it is recognized that a policy is almost never implemented in the way in which it was designed, and in part it depends on the clarity with which the different steps to be followed are specified ( However, we must keep in mind that, in certain experiences, as the Mexican experience very well attests, public programs often begin to be implemented and, *ex post facto*, the design is reflected in guidelines or terms of reference). But, regardless of the thoroughness with which the course is detailed, the implementers will reinterpret and adapt the policy to their environment, which leads us to raise questions such as: how to fight against a customary practice of the Mexican public administration in which. For example, is the procedures manual replaced by the training carried out by the most experienced bureaucrat? In the same way, in his metaphorical study *The Implementation Game*, Bardach (1978) identified the different strategies that bureaucrats implement when they can put a policy into practice, which, thanks to both the monitoring and evaluation of the processes, we will be able to identify and reduce its harmful effects on the policy in question. All of this refers to the so-called implementation gap, which can be identified through the evaluation of processes or operations.

Moving on to the fourth point, the American political scientist refers to the very objectives that the intervention set itself. In his own words: “once evidence about the process and results has been collected, the evaluation determines the merit of the program by comparing the evidence with a series of expectations, and there will always be an element of judgment. Sometimes the criteria applied to make judgments

---

<sup>3</sup> For a better understanding of the internal and external sources of invalidation, it is recommended to consult the texts of Hernández, *et al.*, (2014) & Campbell y Stanley (1973)

comes from the official statement of the program's objectives [...] (Weiss, 2018). At this point, we will have to specify two things: On the one hand, in the first part of the statement we can see that it is necessary for the evaluation to be designed thinking about what the program itself postulates, that is, what is intended? achieve?, which, implicitly, ends up being a criticism of the standardization of evaluations, since we cannot ask that a process carried out for a certain context be used to apply to any program regardless of the objectives<sup>4</sup>.

Although it is true that the establishment of terms of reference, such as those prepared by the National Council for the Evaluation of Social Development Policy (Coneval), can serve as support so that other entities can take advantage of them to evaluate their programs, it is necessary that These terms can be adapted to the specific program. Likewise, it must be considered that, regardless of the normative positions of the evaluator, he must consider the design of the policy and not judge the intervention from its axiological framework, since, if this error is incurred, the program may be completely invalidated, although in fact, he is fulfilling what was expected of him. On the other hand, as we have just seen, Weiss specifies that, regardless of the objective pursued by the evaluation, there is an obligation for the personnel conducting it to make a value judgment to determine to what extent the objectives have been met or not. established by decision makers

The last point that Weiss (2018) refers to is the purpose of the evaluation, and in this case, it must be very clear: the purpose of all evaluation of public policies, and specifically social policies, is: "to contribute to the improvement of the program and policy. Despite this, at this point we must add some things, for example:

In order for the improvement of the program to be achieved, it is necessary that, based on the identified gaps (relevance - that is, that the actions are appropriate to address the identified problem -, congruence - which implies that the objectives can be achieved with the designed means - or implementation - if the processes are carried out as they were designed - (Cardozo, 2012), recommendations are issued that enable their adoption, therefore, only pointing out the failures without indicating the route to follow for their correction can be insufficient to fulfill the purpose of the evaluation. In addition, it is necessary to specify that it would be advisable for the evaluation team to establish the periods that it considers appropriate for its fulfillment, if it considers

---

<sup>4</sup> We could say that, although currently in Mexico, there is a strong tendency to apply the same solution (monetary transfers) to different problems (be it problems of access to basic satisfiers, attention to malnutrition problems, substitution of care services, etc. ), we cannot expect the same type of evaluation to be applied to different contexts, says Cardozo (2012) that Thoenig was concerned since 2002 by "the tendency to apply the same methodological recipes in different contexts, obtaining conclusions from the mere treatment of the data without taking into account an explanatory theory that allows its interpretation"

the time margins to which the evaluations are subject. dependencies, whether for administrative<sup>5</sup> or regulatory reasons<sup>6</sup>.

In this case it is important to take into consideration the debate, precisely between Weiss and Patton about the use of evaluations. Patton, for example, states that since evaluations provide much more than they cost, they must be supported and funded, and that, for him, their use will depend on “knowledge [that] includes the ability to propose an appropriate approach to interact with target users and key stakeholders in order to educate them as information users, and work with them in a mutual commitment to use the evaluation process and its results” (in Delahais, *et al.* 2021: 38) . For his part, Weiss recognizes that “Those responsible for public programs and policies do not systematically use evaluation “as a basis for decision-making” [in addition to the fact that] Mike Patton does not mention the word “policy.” In their world everyone behaves rationally” (Delahais, *et al.* 2021: 54), that is, Patton ignores the weight that politics, or inter- and intra-organizational interests, exert on the evaluation.

In general, the use will depend on various factors, e.g., the existence of a binding regulatory framework (is there a mechanism that forces the implementers to return to the recommendations? - in this case it is worth looking at the existing procedure in Mexico City -<sup>7</sup>), the level of complexity of the recommendation, etc.; In other words,

---

<sup>5</sup> It would not be possible, for example, to issue a recommendation that would imply a budget increase to expand the coverage of the program in the immediate term, without taking into consideration the budget programming deadlines that depend on the fiscal calendars in each administration.

<sup>6</sup> In this case, a recommendation to redesign the operation could not be established without considering the established deadlines in case a reform is required that involves other actors beyond the Executive branch; Let's take as an example what happened recently in Mexico, in which several programs to care for vulnerable groups have been established in the Magna Carta, therefore, promoting a reform to any of these interventions must go through a lobbying process both at the level of commissions and parties in Congress.

<sup>7</sup> Por ejemplo, en la Ciudad de México, la Ley de Desarrollo Social (LDS), establece en su artículo 42D que será el consejo de Evaluación (Evalúa CdMx), quien, con base en los informes finales de las evaluaciones, analizará las recomendaciones y propondrá a su vez las propias, juzgando su viabilidad tanto normativa como política y presupuestaria, acordando un calendario de cumplimiento con el ente evaluado, y en caso de que no suceda de esta manera, se iniciará un proceso de controversia que será dirimido por la Comisión intersecretarial de Desarrollo Social y, “[...] Una vez agotado el procedimiento ante la Comisión el cumplimiento de las recomendaciones será obligatorio, estableciéndose entre el Consejo de Evaluación y el evaluado un programa y calendario para su cumplimiento. La omisión en el cumplimiento de esta obligación será hecha del conocimiento de la Contraloría General del Distrito Federal [sic] y sancionada en términos de la Ley Federal de Responsabilidades de los servidores públicos”. Pese a las aparentes ventajas de este modelo, también se debe considerar que la conformación del calendario y el llamado a la conformación de la Comisión Intersecretarial es una potestad del titular de la Jefatura de Gobierno, es decir, la propia autoridad que instrumenta las acciones de desarrollo social, por lo que lo convierte en juez y parte, lo que puede entorpecer los procedimientos para dirimir las controversias. Por ello, la promulgación de la Constitución Política de la Ciudad de México en el 2017 establece en el artículo 47-3 que, “El Consejo de Evaluación de la Ciudad de México, con base en la ley de la materia, determinará mediante acuerdos generales el número de comités encargados de evaluar respectivamente las políticas, programas y acciones en materia de desarrollo económico, desarrollo social, desarrollo urbano y rural, seguridad ciudadana y medio ambiente. Las recomendaciones que emitan los comités serán vinculantes para orientar el mejoramiento de las políticas, programas y acciones”.

the social context of the evaluation referred to by Thoenig (2018) must be taken into consideration.

In general, the use will depend on various factors, e.g., the existence of a binding regulatory framework (is there a mechanism that forces the implementers to return to the recommendations? - in this case it is worth looking at the existing procedure in Mexico City -), the level of complexity of the recommendation, etc.; In other words, the social context of the evaluation referred to by Thoenig (2018) must be taken into consideration.

In addition to the improvement of the policy, the evaluation can eventually serve for the budgetary control that the Legislative Branch carries out in a regime with a system of checks and balances, this is what O'Donnell calls horizontal accountability. However, in environments such as Mexico, in which budget approval does not require a qualified majority, it is highly feasible that the budget will be defined, practically, in the sense defined by the head of the Executive Branch (if his party has with a relative majority in Congress, as happened for several decades during the PRI hegemony, or as has happened in the last two legislatures in which the President's party has had that advantage in both chambers), that is, make an allocation of resources based predominantly on the interests of the party in power and not on the results identified by the evaluation.

In parallel, evaluation could also be used as a means for vertical accountability (beyond the limited perspective of O'Donnell (1997))<sup>8</sup>. The vertical accountability to which we refer is the one that Schedler (2008) defined by associating it with three components: information, justification and sanction; That is, evaluations would serve as the first component of the Schedlerian triad, providing information to citizens about the effectiveness of policies; in a second moment, citizens could question the authorities about the results obtained and they would have the obligation to argue, substantiate and motivate the reasons for said results; and finally, if the response is not satisfactory or justified, we could reach the third and final component: that of the application of sanctions, a situation that, however, must be qualified, something that we will do in the next paragraph.

Within the evaluation we can find two approaches: the punitive and the formative. Punitive is one in which the evaluation serves to apply sanctions, whether of a political-electoral, budgetary, administrative or criminal nature, against public authorities or officials, either for failure to comply with their obligations or because the objectives are not achieved. expected results; This will cause, in part, reluctance

---

<sup>8</sup> Horizontal accountability, according to O'Donnell (1997), is what is carried out when power monitors power, that is, when there is an effective system of checks and balances between the Executive, Legislative and Judicial powers, in so much so that vertical accountability is identified, almost exclusively by the Argentine political scientist, with electoral processes, where citizens have the power (from time to time) to reward or sanction the authorities.

and distrust towards the evaluation, precisely because we think about the negative consequences that it brings with it. On the other hand, the training or learning approach, according to Nioche, represents the search for the evaluator to collect and judge the evidence, and serve as “facilitator of the learning process, which implies a change in role and characteristics.” required to fulfill it (interdisciplinarity, negotiation skills, didactics, etc.)” (cited in Cardozo and Mundo, 2012: 28).

In Mexico, since the late 1990s, performance evaluation was introduced through the postulates of the NPM, and it had a strong punitive nature, that is, punishing officials who did not achieve the objectives entrusted to them. However, for some years now, both Coneval and Evalúa, to give an example, have been adopting a perspective that adopts more for learning than for sanction. However, we must make it clear that policies that do not obtain adequate results must be modified or replaced since these are interventions that fail to establish the expected changes in the population they serve.

In addition to the above, a couple of additional issues must be considered. Firstly, if the evaluation has a single purpose (the improvement of public interventions, whether called policies, programs or projects) and different uses (horizontal accountability through budget control, democratic exercise, vertical accountability, etc.), It must be kept in mind that evaluation is not an end. If for a moment we adopt Patton’s romantic vision, where it seems that the information provided by the evaluation serves, yes or yes, for the improvement of policies independently of the political and technical processes, it is because the evaluation is part of a whole process, the process of all public policy and, therefore, we must consider the evaluation within an entire planning scheme.

Yehezkel (1980), during the last decades of the last century, proposed the establishment of a general planning model in which, from the diagnostic phase, it was necessary to think about carrying out evaluation activities throughout all the stages of the policies, therefore, ex ante, ex post and ex tempore evaluations had to accompany both the formulation and implementation of the policy. Also, Cohen & Franco (1992) established a general framework in which evaluation was not separated from the rest of the activities of public programs, but rather went hand in hand, serving as a feedback process.

The word planning began to suffer from a stigma that associated it, in the context of the Cold War, with the socialist bloc (due to the five-year planning carried out in the former Union of Soviet Socialist Republics). Luis Aguilar says that this type of planning was associated with the exhaustive rational model<sup>9</sup>, typical of totalitarian societies, which implied an omnicomprehensive view, as well as a high capacity for

---

<sup>9</sup> A la letra, Aguilar menciona lo siguiente: “En un frente los racionalistas, analistas sinópticos, comprensivos, planificadores, maximizadores...; en el otro frente, los incrementalistas, racionalistas limitados, pluralistas, los politólogos de las [ciencias de las] políticas” (2000: 59)



processing and analyzing information that allowed us to know not only the causes, but also the consequences of each alternative action (2000, 2000a). Later, the adoption of the New Public Management model, which sought to establish the private management administration model in government organizations without considering the enormous differences between both fields, continued in its concealment of the importance of planning, and has not been but until very recently when the importance of betting on planning has been reaffirmed.

Finally, in this regard, we must say that, without attempting to make a tiring summary of the different typologies that exist around evaluation (ex ante, ex post, ex tempore, internal, external, mixed, participatory, design or formulation, of processes or implementation, of results, of impact), since Cardozo & Mundo (2012), Cardozo (2012), Franco & Cohen (1992), have already paid To this objective, it is necessary to say that we must differentiate between evaluation and evaluative research, since the latter is only concerned with the accumulation of knowledge, while the former has “applied and instrumental” purposes (Cardozo & Mundo, 2012), and although Weiss refers interchangeably to evaluation and evaluative research, the former is characterized by its practical nature, while the latter is content with making findings; The first is carried out at the request of the implementers who seek to improve public action, while the second is initiated by a personal initiative on the part of researchers, generally associated with university research centers. In short, only evaluative research represents an end, while evaluation broadens its horizons. Evaluation has value only to the extent that it allows for the redesign of policies and the improvement of decision making.

The second aspect to consider, and with which we intend to conclude this work, is closely related to the previous one. If we assume again that some sectors consider evaluation as an end, without considering that the usefulness of evaluation goes beyond identifying the results of the policy, in recent years there has also been a metonymic process, in which that a part assumes the position of a whole, in this case we are referring to the reification that the word evidence has experienced. In recent years, a current of thought in which we can mention the texts of Head (2010) and Merino *et al.* (2021), as well as the holding of multiple seminars, such as that of the Instituto 512 (2022) or that of Conevalvideo (2023), we see that evidence is put before (and not evaluation) as the basis for decision-making.

At the outset, we must say that the word evidence is a copy of the English evidence, and the Pan-Hispanic Dictionary of the Spanish Language, indicates the translation as accurate, both in its quality of concrete manifestation and abstract quality, but identifies an enormous abuse in the use of this word, and adds that it is not justified:

*“The indiscriminate use in Spanish of the word evident as a synonym for proof or indication, a reprehensible copy of the English evidence: in English, evidence is any evidence (circumstantial, testimonial, material, documentary, etc.) that is alleged in a judicial process; In Spanish, it would only be acceptable as a synonym for evident proof, that is, clear and manifest proof; Thus, uses such as the following are not appropriate: “The evidence that has been provided does not seem very convincing in any case» (Ninyoles Idioms [Esp. 1977]);” “The circumstances and evidence were clear against Dr. Sittón” (Siglo [Pan.] 12.5.1997)” (RAE)”.*

On the other hand, Hausmann (2016) mentions that, “it is sensible to demand that policies be based on evidence and that this be the best possible, within reasonable limits of time and budget. However, the way this approach is now implemented may be causing a lot of harm by weakening our ability to learn and improve what we do.” This former Venezuelan planning minister warns that this trend comes, as we said previously, from the preponderance that exists in the application of experimental methods (with their epitome, randomized control trials), however, this preponderance of evidence can lead to attribution errors, such as thinking that the application of a certain method (such as the introduction of a new technology to improve academic achievement - that is, evidence) can be extrapolated to other environments in which It will not yield the same fruits. Therefore, Hausmann advocates a process of learning and concomitant feedback.

On the other hand, it is necessary to detail that, in legal language, a criminal is not sentenced based on “evidence”, but rather this serves to contextualize the way in which a certain crime is committed. The evaluation of this evidence (or evidence) by the judge and jury is what will determine the guilt or innocence of the accused. Transferring this reasoning to evaluation, it is not the evidence that allows the improvement of a certain policy, it is the entire evaluation process in which they play an important role (but not the only one), since they allow certain statements to be made about politics, but they will not be the only factor that allows the value judgment to be made. Evidence is part of a whole process, so it cannot take the form of that whole, since, as Hausmann says, this evidence can be judged wrongly and may not produce the same results in different contexts.

## CONCLUSIONS

Herein, the definition of the evaluation activity was specified, not in its colloquial sense, as the cited authors have helped us identify, but in its manifestation as part of the policy process, therefore, evaluation has its It was valuable to the extent that it fulfills its main purpose (improvement in decision-making for the redesign of politics), and can help other types of activities carried out in a democratic regime.

Based on this, it is concluded that policies, programs and projects can improve their results to the extent that they are evaluated, and despite its weaknesses, a binding system can better help to get the most out of them. this type of efforts (despite the various limitations it presents); although we cannot ignore a system like the one that Coneval implements with its Follow-up to Recommendations of External Evaluations, which implicitly assumes that officials and decision makers will do what is most convenient for public interventions.

For all these reasons, evaluation is an indispensable tool to expand the exercise of political and citizen rights, in a democratic regime, but also to pay for the fulfillment of economic, social, cultural and environmental rights.

#### LITERATURE CITED

- Aguilar, L. (2000). *The study of policies*. Miguel Ángel Porrúa.
- Aguilar, L. (2000a). *The making of policies*. Miguel Ángel Porrúa.
- Bamberger, M., Rao, V., & Woolcock, M. (2018). Use of mixed methods in monitoring and evaluation: Experiences in international development. In G. Pérez & C. Maldonado (Eds.), *Anthology on impact evaluation* (pp. 380-402). CIDE.
- Bardach, E. (1978). *The implementation game*. The MIT Press.
- Campbell, D., & Stanley, J. (1973). *Experimental and quasi-experimental designs for research*. Amorrortu.
- Cardozo, M. (2012). *Evaluation and meta-evaluation in public policies and programs: State of the art*. UAM.
- Cardozo, M., & Mundo, Á. (2012). *Guidance manual for the evaluation of social development policies and programs*. Evalúa-DF.
- Cohen, E., & Franco, R. (1992). *Evaluation of social projects*. Siglo XXI.
- Conevalvideo. (2023, [exact date]). *CONEVAL-INAFED Forum: Evidence for decision-making in Mexican municipalities* [Video]. YouTube. <https://www.youtube.com/watch?v=YuknMwf7P6Y>
- Deleau, M., Nioche, J. P., Penz, P., & Poinard, R. (1986). *Evaluating public policies: Methods, ethics, organization, report of the working group*. La Documentation Française.
- Delahais, T., Devaux-Spatarakis, A., Revillard, A., & Ridde, V. (2021). *Evaluation: Foundations, controversies, perspectives*. Éditions Science et Bien Commun.
- Dror, Y. (1980). *A general planning model*. DCPyS, UNAM.
- Hausmann, R. (2016, February 25). The problem with evidence-based policies. *Project Syndicate*. <https://www.project-syndicate.org/commentary/evidence-based-policy-problems-by-ricardo-hausmann-2016-02/spanish>
- Head, B. (2010). Evidence-based policy: Principles and requirements. In *Strengthening evidence-based policy in the Australian federation* (pp. 13-26). Productivity Commission.
- Instituto 512. (2022, [exact date]). *Webinar: Evidence-based decision making* [Video]. YouTube. <https://www.youtube.com/watch?v=f8vQtI7OVOE>
- Hernández, R. (2014). *Research methodology*. McGraw-Hill.
- Meny, Y., & Thoenig, J. C. (1992). *Public policies*. Ariel.
- Merino, M. (2021). *Accountability in emergency situations: An evidence-based model*. Universidad de Guadalajara.
- O'Donnell, G. (1997). Horizontal accountability and new polyarchies. *Nueva Sociedad*, 152, November-December, 143-167.

- Pressman, J., & Wildavsky, A. (1998). *Implementation*. Fondo de Cultura Económica.
- Real Academia Española. (2005). *Pan-Hispanic dictionary of doubts*. Espasa Calpe.
- Schedler, A. (2008). What is RC? IFAI. [http://www.iepcjalisco.org.mx/sites/default/files/que\\_es\\_la\\_rendicion\\_de\\_cuentas.pdf](http://www.iepcjalisco.org.mx/sites/default/files/que_es_la_rendicion_de_cuentas.pdf)
- Thoenig, J. C. (2018). Evaluation as actionable knowledge for public management reforms. In C. Maldonado & G. Pérez (Eds.), *Anthology on evaluation* (pp. 269-292). CIDE.
- Weiss, C. (2018). Preparing the ground. In C. Maldonado & G. Pérez (Eds.), *Anthology on evaluation* (pp. 43-84). CIDE.

The logo for REMEVAL, featuring the word "REMEVAL" in a blue, sans-serif font. The letter "e" is stylized with a yellow and orange gradient and a curved underline.