



HAL
open science

Social impact measurement and public management: how social innovation challenges the policy evaluation paradigm –the case of France

Bryan Dufour

► **To cite this version:**

Bryan Dufour. Social impact measurement and public management: how social innovation challenges the policy evaluation paradigm –the case of France. 8th International Social Innovation Research Conference: "Social Innovation in the 21st Century: Beyond Welfare Capitalism?", Sep 2016, Glasgow, United Kingdom. hal-01458735

HAL Id: hal-01458735

<https://hal.science/hal-01458735v1>

Submitted on 6 Feb 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Social impact measurement and public management: how social innovation challenges the policy evaluation paradigm –the case of France

8th ISIRC Conference

September 5-7 2016

Social innovation in the 21st century: Beyond welfare capitalism?

Stream: Public policy and social innovation: comparative evidence, critical assessments

Bryan Dufour (*)

bryan.dufour@gmail.com

(*) PhD fellow. Aix Marseille Univ, CNRS, LEST, Aix-en-Provence, France

Abstract: As an increasing number of governments draw an interest in social enterprises and seek to understand the impact of social innovation through social impact measurement (SIM) practices, this paper builds a gap analysis between “traditional” programme evaluation (PE) as it is carried out by public actors and SIM as it is practiced by social enterprises. After framing our contribution in terms of context and definitions, we proceed with the case study of France, where we compare public praxis for both SIM and PE based on a documentary analysis. We find that both disciplines are bridged by a common theoretical foundation and, to a certain extent, by participative approaches. We also identify three main gaps, which are (i) the way outcomes are treated in PE and SIM; (ii) how the stakeholders’ participation is managed and how it affects the ownership of the evaluation process by the involved parties; and (iii) how metrics and indicators are approached. This paper is part of a broader research project on SIM focused on work integration social enterprises (WISEs) in the French and Danish public management contexts. The SIM and PE approaches studied here are therefore considered in the perspective of an application for WISEs.

1. Introduction: frame and context of this paper

1.1. Context: the place of social enterprises in welfare capitalism

1.1.1. Public management reforms and the push for evaluation

The development of public management reforms in the last decades, largely influenced by new public management (NPM, Hood 1991) has driven a growing attention on programme evaluation (PE, Dunn & Miller, 2007; Lapsley, 2009). In its quest for introducing management apparatuses inspired from the private sector into public administration, the NPM paradigm argues that evaluation provides the information policy makers need to make better-informed choices, ultimately improving the efficiency of the public sector. Also stemming from NPM, the spurring of public-private partnerships (PPPs,

Broadbent and Laughlin 2003; Edwards and Shaoul 2003; Grimsey and Lewis 2007) have fostered new ways of delivering goods and services (including welfare-related services) traditionally channelled by the public sector. This trend also contributed to the development of a culture of evaluation within public management, facing the need to assess a growing number of third parties involved in the delivery of public goods.

1.1.2. Public management reforms and the place of social enterprises in the new welfare state

Several contributions specifically emphasize how public-private partnerships involving social enterprises can contribute to an increased welfare (Albareda et al., 2007; Evers & Laville, 2004). Furthermore, within the NPM paradigm, social enterprises can appear as a particularly appealing resource from the private sector to leverage in the public sphere for governments seeking potential ways to increase the efficiency of the services to deliver (Hulgård, 2011; A. Nicholls, 2010; Nyssens & Defourny, 2012; Pestoff, 2008). Just like PPPs in general, the development of social enterprises found a positive echo in the public sphere through the political reforms enthused by NPM, where they appeared as an additional device to inject further market logics into public administration. Social enterprises are therefore progressively being recommended as a key lever for governments to focus on in order to alleviate part of the economic and social challenges they encounter (European Commission & OECD, 2013; Hulgård, 2011).

1.1.3. Social enterprises and impact measurement: another form of evaluation?

This enthusiasm for social enterprises, as well as the development of impact investing practices, has led to an increasing focus on social impact measurement (SIM, Emerson, 2003). This trend has first emerged among private actors, eager to prove and manage their impact. While the development of impact measurement has been accelerating in the past few years, a growing number of specialised

organisations have emerged (e.g. B Lab, Social Asset Measurements or Sinzer), as well an ever-expanding list of methodologies for SIM (Grieco, 2015; Olsen & Galimidi, 2008).

Public stakeholders have also picked up this infatuation for impact measurement, which happens to be a well-fitted feature for the performance-driven NPM and its focus on evaluation. In this context, we observe an increasing number of attempts to make SIM part of regulatory frameworks aiming to favour the third sector for the provision of certain welfare services (e.g. in 2014 the G8's Social Impact Investment Taskforce or the European Commission's SIM Sub-group -GECES).

Our initial position is therefore to suggest that SIM, as it is currently recommended by policy frameworks, is a form of PE where the programme to assess is a social enterprise. This brings us to our question of understanding whether SIM, as an offspring of social innovation, has the capacity to alter the century-old practice of PE (Hogan, 2010; Scriven, 1996). While not bringing a definitive answer to this question, this paper proposes to understand what may differentiate SIM from PE investigating the case of France, looking at how both disciplines are promoted and executed in the public spheres, from a WISE perspective.

1.2. Framing our approach: definitions

1.2.1. Programme Evaluation

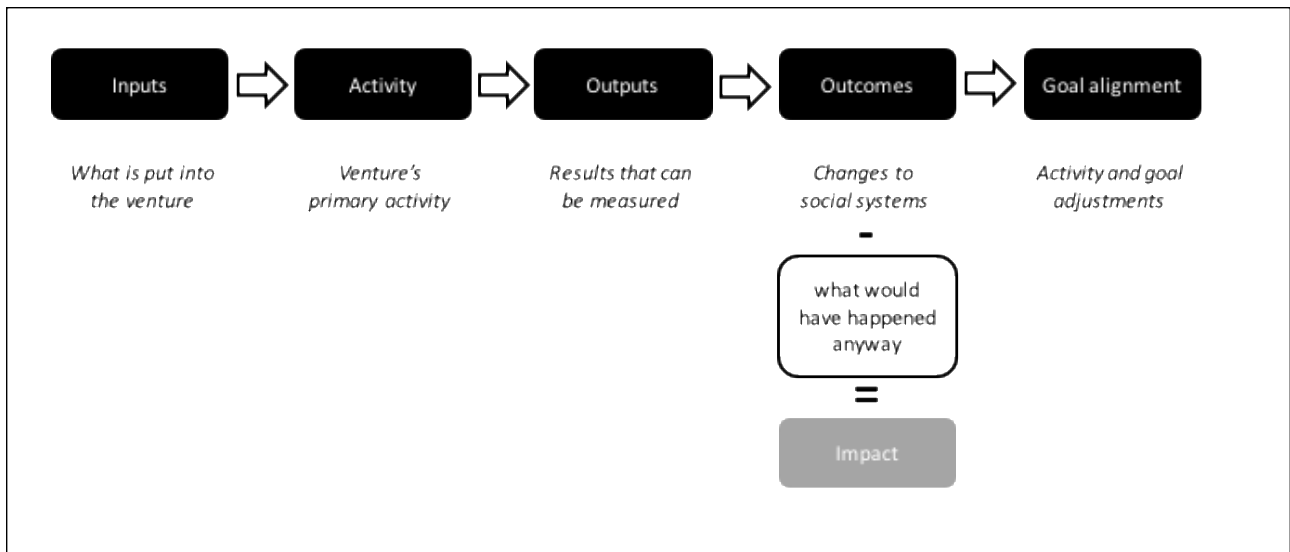
Key authors in our review of the PE literature include Borus, 1979; Furubo, Rist, & Sandahl, 2002; Heckman, LaLonde, & Smith, 1999; Imbens & Wooldridge, 2009; Levitan & Wurzburg, 1979; Poister et al., 1981; Rubin, 1973; Scriven, 1991 and Stufflebeam, 2001.

As highlighted by Pollitt (1993, p.353), “There are many slightly varying definitions of evaluation”. Scriven (1991, p.139) provides a concise definition: “judging the worth or merit of something or the product of the process”. The aspect of “merit or worth” appears central in several other definitions (e.g. Guskey, 1999, p.41; Stufflebeam, 2001, p.11) and comes from the Joint Committee on Standards for Educational Evaluation (JCSEE)’s *Program Evaluation Standards* (Yarbrough, Shulha, Hopson, & Caruthers, 2010). While the field of application of such PEs can be limitless, and because of our interest in work integration social enterprises (WISEs), we will focus on PE methods applicable to initiatives targeting social issues such as unemployment and work integration. This paper therefore adopts the following definition: “Program evaluation is the use of social research methods to systematically investigate the effectiveness of social intervention programs in ways that are adapted to their political and organizational environments and are designed to inform social action to improve social conditions.” (Rossi, Lipsey, & Freeman 2003, p.16).

1.2.2. Social impact measurement

Furthermore, as we are concerned with SIM we will adopt the following definition for what constitutes social impact: “Social impacts include all social and cultural consequences to human populations of any public or private actions that alter the ways in which people live, work, play, relate to one another, organize to meet their needs, and generally cope as members of society” (Burdge & Vanclay, 1995, 59). The cornerstone of most of the SIM methods developed in the last decades is the impact value chain (Clark, Rosenzweig, Long, & Olsen, 2004), as depicted in Figure 1. While we retain this specific definition, we acknowledge the debate in academia and among practitioners around the definition of terms such as “social value”, “social impact”, “social value creation” or “social return” (Austin, Stevenson, & Wei-Skillern, 2006; Gibbon & Dey, 2011; Maas & Liket, 2011; J. Nicholls, 2007).

Figure 1: impact value chain



Adapted from Clark et al., 2004

1.2.3. Work integration social enterprises

Finally, and as this paper is part of a broader project on work integration social enterprises (WISEs), our research focuses on this specific subset of social enterprises. While the definition of what is a social enterprise remains debated among both academia and practitioners (Dees et al., 1999, Ewing, Haas, & Haas, 1998; Mair & Martí, 2006; Nyssens & Defourny, 2012; Yunus, 2006), the concept of WISE offers the substantial benefit of being relatively well and consensually framed. Davister, Defourny, & Grégoire provide the following definition: “WISEs are autonomous economic entities whose main objective is the professional integration – within the WISE itself or in mainstream enterprises – of people experiencing serious difficulties in the labour market. This integration is achieved through productive activity and tailored follow-up, or through training to qualify the workers.” (2004).

2. Programme evaluation and social impact measurement: the case of France

2.1 A short history of programme evaluation in France

The first attempt to introduce formal evaluation mechanisms in the French policy making process was the “rationalisation of budgetary choices”¹ introduced in 1968 and inspired by the American Planning Programming Budgeting System (Perret, 2006). The scheme, facing many criticisms, was abandoned in 1983. Since then, several evaluation initiatives took place, chiefs among them are the creation of the Scientific Council for Evaluation² in 1990, changed into the National Council for Evaluation³ in 1998 and the passing of the LOLF⁴ in 2001, an organic law reframing policy funding. Following this, the reforms launched by the General Review of Public Policies⁵ (RGPP, 2007 – 2012) accelerated the need for programme evaluation.

This was acted in the 2008 constitutional reform, which marked the institutionalisation of evaluation practices in the French constitution (Fouquet, 2014). Most notably, the modified version of Article 24 specifies that the Parliament has a formal role of evaluation: “Parliament shall pass statutes. It shall monitor the action of the Government. It shall assess public policies”⁶ (Assemblée Nationale, 2016). The last phrase was added in 2008 and the French “évaluation des politiques publiques” (policy evaluation) is usually equated to “programme evaluation” in the international literature. The reform further ascribes a formal role of evaluator to the Court of Auditors⁷ in article 47-2: “The Cour

¹ Rationalisation des Choix Budgétaires

² Conseil scientifique de levaluation (CSE)

³ Conseil national de l'évaluation

⁴ Loi organique relative aux lois de finance (LOLF)

⁵ Révision générale des politiques publiques (RGPP)

⁶ Le Parlement vote la loi. Il contrôle l'action du Gouvernement. Il évalue les politiques publiques.

⁷ Cour des comptes

des Comptes shall assist Parliament in monitoring Government action [...] in assessing public policies”⁸ (Assemblée Nationale, 2016).

2.2 Current evaluation praxis in France: evaluation in the context of the MAP

In 2012, the Modernising of public action (MAP)⁹ took over the RGPP. One key feature of the MAP is that it provides a clearly framed methodology for the way evaluations should be carried out. It establishes a secretariat (SGMAP)¹⁰ under the direct supervision of the Prime Minister that, amongst other, provides guidance and a streamlined process for evaluation. The SGMAP defines 4 core principles guiding policy evaluation in France (SGMAP, 2015):

1. Decision support: evaluations’ output should be actionable information in order to take the best possible decisions to support policy development.
2. Participation and transparency: all stakeholders should be involved in the evaluation process, and results should be communicated in a transparent manner.
3. Multi-criteria analysis: The evaluation should encompass the entirety of the programme’s value chain
4. Methodology: Apply the three first principles in a frame that is recognised by all stakeholders, including the evaluation’s owner¹¹, the program’s beneficiaries and the evaluators.

⁸ La Cour des comptes assiste le Parlement dans le contrôle de l’action du gouvernement [...] dans l’évaluation des politiques publiques

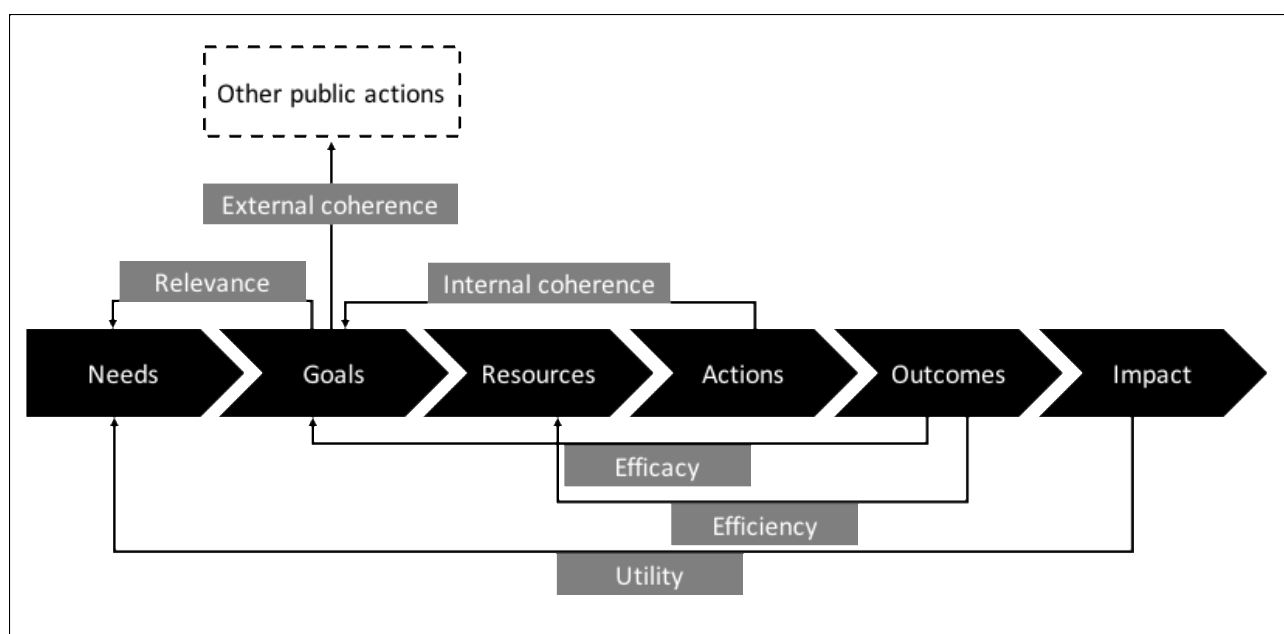
⁹ Modernisation de l’action publique (MAP)

¹⁰ Secrétariat général pour la modernisation de l’action publique (SGMAP)

¹¹ Commanditaire / maître d’ouvrage

The SGMAP also provides a “public action value chain”¹² (cf. Figure 2), quite reminiscent of the SIM’s social value chain and of other performance frameworks associated with public management reforms (see for instance Pollitt & Bouckaert, 2011, pp. 16, 133). The available framing document from the SGMAP also provides a list of performance criteria attached to the value chain (in grey in Figure 2), but does not describe them. This is around these criteria that the evaluation’s owner is supposed to frame the questions they want to be answered.

Figure 2: Public action value chain as presented by the SGMAP



SGMAP, 2015, p.10

The last key aspect of the methodology provided by the SGMAP is the breakdown of the evaluation process in three key stages:

¹² Chaîne de valeur d’une action publique

1. Operational framing: it is the early stage where evaluative questions are defined and the overall approach is agreed upon by the main stakeholders involved in the evaluation.
2. Diagnostic: the research and analysis phase where the current situation is assessed and the evaluation questions answered.
3. Scenarios: the last phase where, drawing from the diagnostic, the evaluators draft different transformative scenarios for the considered programme and estimate impacts for each of the considered scenarios.

After these three steps follow the publication of the evaluation report and the decisions taken on the basis of its conclusions.

2.3. The MAP in practice

Transparency is one of the core principles promoted by the SGMAP regarding PE in the context of the MAP. This vision was particularly well executed through a website regrouping all the PEs carried out since the start of the MAP in 2012, gathering all key documents pertaining to each evaluation (engagement letter, evaluation report and other ad'hoc reports). We used this website as a primary source of information to get an understanding of how PEs are actually carried out under the MAP.

At the time of the research¹³, the online repository provided access to 78 evaluations, of which 50 had been completed, with a published evaluation report. We used our focus on WISEs to build a sample of five evaluations. We selected programmes which activities could potentially be associated to a WISE (such as social inclusion, professional training and insertion).

¹³ Updated in August 2016

Table 1 provides an overview of the five selected evaluations and their state of advancement at the time the research was done.

Table 1: Selection of relevant MAP programme evaluations

Evaluated programme	Start date	Evaluation closing date	Last stage executed	Key documents available
Territorial governance for inclusion policies	March 2013	December 2013	Decision	Letter of engagement, evaluation report
Insertion of incarcerated individuals	May 2015	N/A	Report (not published)	Letter of engagement
Offer and demand of social and socio-professional support	December 2015	N/A	Operational framing	Letter of engagement
Vocational training for unemployed individuals	February 2013	August 2013	Decision	Letter of engagement, evaluation report
Partnership between the national education institutions and the business sector for the professional inclusion of the youth	October 2014	October 2015	Published report	Letter of engagement, evaluation report

While evaluation reports are not always available for the selected PEs, the letters of engagement are for all of them. They provide the concrete evaluation questions evaluators are asked to answer. These letters have therefore been used as a foundation for our analysis to assess the general direction of the commissioned evaluation. They allow us to understand the evaluation's priorities and the angle under which these priorities are approached. The engagement letters nonetheless do not allow to get a proper understanding of which research tools and analytical methodologies are applied (e.g. stakeholder interviews, focus groups, statistical surveys, economic estimates, etc.). However, the three evaluation reports available for the selected PEs, as well as more reports selected at random among other evaluations reveal that in most cases, evaluations are carried out through a combination of quantitative and qualitative tools, with an emphasis on the latter.

2.4. Social impact measurement in France

In the past few years, France has started to formalise an institutional network dedicated to the development of the “social and solidary economy”¹⁴, of which social enterprises are a part. This arrangement includes, among others, a High Council on Social and Solidary Economy (CSESS)¹⁵, a French Chamber of Social and Solidary Economy with a network of regional chambers and a dedicated Observatory. This was all accelerated and strengthened by the act “on social and solidary economy” passed in July 2014 and providing a legislative grounding for the stakeholders mentioned above as well as a legal framework for fostering social enterprises in the country.

This growing interest for social enterprises has also fostered an awareness for SIM. Coined as social impact measurement¹⁶, the concept appears as a fairly new item in the political agenda, with a first milestone in 2011 through a dedicated report commissioned by the CSESS. There was however an existing discussion about the impact of the third sector in France prior to this report. A growing number of stakeholders were then preoccupied with measuring “social utility”, through a trend started in the early 2000s¹⁷ (e.g. Dauphin, 2012; Duclos, 2007). To a large degree, SIM should therefore be considered a prolongation of the already existing reflection on the measurement of “social utility” in France. To this extent, we parenthetically highlight that the report commissioned by the CSESS uses both “social impact” and “social utility” interchangeably in two instances (CSESS, 2011, pp. 14;20).

Although public stakeholders such as the French Chamber of Social and Solidary Economy touches upon the theme of SIM, there are no formal guidelines emanating from a public body regarding impact

¹⁴ Economie sociale et solidaire (ESS)

¹⁵ Conseil Supérieur de l'Economie Sociale et Solidaire (CSESS)

¹⁶ Mesure de l'impact social

¹⁷ Mesure de l'utilité sociale

measurement practices or methods. However, AVISE¹⁸, an association created by the French Deposits and Consignments Fund¹⁹ and several other stakeholders from the third sector, offers in-depth documentation in impact measurement, covering methodologies, processes and stakes. On top of its genesis anchored in public institutions with the Deposits and Consignments Fund as a founding member, AVISE acts as an Intermediate Body for the European Social Fund and benefits from the sponsoring of several ministries. It therefore has a strong legitimacy and is often associated to a public body by the stakeholders it interacts with.

We will thus take AVISE's guidelines regarding SIM as the closest it gets to state-sponsored recommendations in France. AVISE offers a rather wide collection of documents regarding SIM: it has edited and hosts 25 web pages dedicated to SIM, and references 14 reports from a variety of stakeholders (mostly academics, governmental bodies, international institutions, foundations and consulting companies). We highlight that although the SIM methods promoted by AVISE are not specifically targeting WISEs, all of them can be applied to gauge the kind of impacts outlined by the definition of a WISE. Furthermore, a substantial share of the cases put forward by AVISE in its documentation, where SIM have been applied, are initiatives that fall into the definition of a WISE (e.g. Ares, MillRace IT, Acta Vista, etc.). Table 2 provides an overview of these reports, specifying whether they propose a selection of SIM methodologies and / or provide at least one illustrative case.

¹⁸ Agence de valorisation des initiatives socio-économiques

¹⁹ Caisse des dépôts et consignations

Table 2: Main social impact measurement-related reports referenced by AVISE

Title (original language)	Author	Year of publication	Content:	
			SIM methods	Case
Petit précis de l'évaluation de l'impact social	AVISE, ESSEC	2013	Yes	Yes
La mesure de l'impact social	CSESS	2011	Yes	No
Approches proposées pour la mesure de l'impact social	EU /GECES	2014	No	Yes
Analyser, suivre et évaluer sa contribution au changement social	AFD	2016	No	Yes
Measuring Impact	G8	2014	No	Yes
Guide de la mesure d'impact social	Rexel Fondation, Improve	2014	No	No
Guide du retour social sur investissement (SROI)	ESSEC	2011	No	Yes
Un guide pratique pour la mesure et la gestion de l'impact	EVPA	2015	No	Yes
Évaluer l'utilité sociale de son activité	AVISE	2007	No	No
Passeport Avenir, un programme qui compte double	Accenture / ESSEC	2012	No	Yes
Évaluer l'impact social d'une entreprise sociale : points de repère	Stievenart and Pache	2014	Yes	No
Évaluer l'utilité sociale de l'économie sociale et solidaire	Branger, Gardin, Jany-Catrice, Pinaud	2014	Yes	Yes
Faut-il tout mesurer? Réflexions sur l'utilité de la mesure de l'impact social	Mortier	2013	No	No
Guide de bonnes pratiques	ESF	2015	No	Yes

3. Differences and commonalities between SIM and PE in France

The gap analysis presented below was created by going in a systematic way through the documentation made publically available by AVISE (for SIM practices in France) and the SGMAP on the five selected cases presented in

Table 1 (for PE practices in France). From these documentary resources, we infer what are presented as recommended practices in France for both SIM and PE. With the evaluation process made public in the context of the MAP, our study of PE benefits from more depth, allowing to compare the

methodological frame promoted by the SGMAP and the way it is put in practice by evaluators and evaluations' owners. This is unfortunately not the case for SIM where the only cases available are the ones selected by the stakeholders whom AVISE is relaying the publications. We therefore choose not to exploit these cases for analytical purpose (i.e. their content has not been qualitatively coded) as they may reflect obvious instances of biased selectivity (Yin, 2009, p. 102). We however use them to inform our understanding of some of the SIM methods promoted by AVISE.

The documents we used for SIM are therefore the content edited, hosted or shared by AVISE on its website, as presented in page 12 of this paper. For PEs, we used the general documentation provided by the SGMAP to communicate about its methodology to internal and external stakeholders (policy reports, memos, guidebooks and templates) as well as the material available from the five PE cases we selected (cf. Table 1): five engagement letters and three evaluation reports, complemented by ad'hoc research on the evaluated programmes. We carried out the documentary analysis by coding the information contained in AVISE and the SGMAP documents around key words and concepts, following a constant comparative method (Glaser & Strauss, 1967) and identifying patterns, clusters and differences across the documentary corpus (Bowen, 2009).

3.1. Easier said than done? Drawing a line between programme evaluation and social impact measurement

Looking at France's case, our initial stance of assuming that SIM may be PE in disguise remains, to some extent. We identify two main commonalities between the two disciplines.

3.1.1. A common theoretical foundation

Both PE and SIM are underpinned by very similar theoretical tenants in most of their respective literature: symbolizing this alignment is the use of the “public action value chain” for the former (cf. Figure 2, also called “logic model” in most of PE contributions) and the impact value chain for the latter (cf. Figure 1). Both disciplines accept the same key concepts and arrange them in very comparable fashions to reach similar ends: the understanding of outcomes and impact, and the processes leading to them. Resemblance in the taxonomy is even pushed further, with the “programme theory” in PE (Rossi et al., 2003) and the “theory of change” in SIM (J. Nicholls, Cupitt, & Durie, 2009). On top of this common theoretical frame, both PE and SIM use the same analytical and research tools (interviews, surveys, economic and financial estimates, statistical measurements, etc.) to gauge and understand impact. All these observations are strongly evidenced in the methodology promoted by AVISE and in the procedures as well as the engagement letters and evaluation reports made available by the SGMAP.

Distancing ourselves from the specifics of the French case, and looking at historical developments, it is fair to say that PE has largely informed the development of SIM, even though some SIM methodologies also take cues from other discipline (e.g. the influence of accounting in the development of frameworks such as AA1000 or SA8000²⁰). The genesis of SIM in the 1990s shows a clear anchoring into the PE approaches developed in the late 1960s (e.g. social impact assessment studies) and imposed by governmental regulation -primarily in the USA. This explains in large part the common theoretical framework we highlight here and puts perspective on how SIM, as a practice

²⁰ Two “social accountability” frameworks, created respectively by AccountAbility and Social Accountability International.

spun out of PE by private actors to develop new ways of evaluation, is being looped back into the political agenda as a substitute for PE.

3.1.2. Stakeholders participation

The participative nature of the PE framework proposed in France is another common foundation with most SIM practices. Stakeholders involvement is very pregnant in SIM methods and is at the centre of the recommendations advocated by AVISE. In the French PE environment, this aspect is also strongly emphasised by the official SGMAP guidelines, where it is anchored as one of the four core principles. Raising participation at this level of the PE framework in France is a novelty introduced by the MAP and allows, at least in theory, to move evaluation away from a traditional top-down approach towards a more inclusive (yet still not bottom-up, cf. below) process where stakeholders can inform the evaluation process and take a more active part in it. Given the recent adjunction of this aspect to the French PE praxis, one could argue it takes cues from the recent developments of SIM methods. The evaluation guide initially published during the inception of the MAP (Battesti, Bondaz, Marigeaud, & Destais, 2012), while strongly stressing the necessity of a participative nature for PE, seems to indicate otherwise: it highlights the main references having influenced the proposed approach, but all of them are related to public policy evaluation and none to SIM (Battesti et al., 2012, pp. 35-39).

3.2. Mind the gap

Despite these conceptual similarities, there is a number of fundamental differences in the way SIM and PE are carried out at the operational level. These differences are the foundation of the gap analysis we are building in this paper and are described below.

3.2.1. Ownership: top-down or bottom-up?

One key differentiation between PE and SIM appears in the ownership of the evaluation process. In the MAP framework, the Prime Minister is the one framing each PE before operational details are trickled down to relevant ministries, local administration (altogether project owners¹¹) and appointed evaluators (SGMAP, 2015). Having the head of government so closely woven into this process can surely be seen as the pinnacle of top-bottom approaches. By contrast, SIM practices often offer a much more organic and bottom-up attitude where all stakeholders are invited from the onset to define the orientation of the evaluation to be carried out. Stakeholders' involvement throughout the SIM process (i.e. even in the preliminary stages) is a key highlight that stands out from the documentation provided by AVISE (AVISE, ESSEC, & Mouves, 2013; G8's Social Impact Measurement Taskforce, 2014; Rexel Fondation & Improve, 2014). It is aimed at drastically impacting the ownership of the entire SIM process, involving all stakeholders and promoting a form of common sponsorship of the SIM and its outcomes (CSESS, 2011; ESSEC, 2011; European Commission / GECES, 2014).

We recall however, as highlighted above, that the French PE paradigm as presented by the SGMAP formally highlights that evaluation should be participative and include all the programme's stakeholders. To this extent, it is undeniable that such participation occurs in practice, as evidenced by the available PE reports (e.g. the 250 interviews with local programme managers and the focus groups with programme users documented in the PE's report for the territorial governance of inclusion policies). However, overall stakeholder participation happens mostly during the research stage and to a limited extent before. This substantially impacts the ownership dynamics of the PE, especially for programme's beneficiaries, who cannot be involved in the early stages of the evaluation process.

It is understood that due to obvious scale considerations, PE implementation cannot mirror SIM practices in terms of stakeholders' involvement. However, some countries like Norway manage to considerably increase the level of stakeholders involvement in their PEs (Furubo et al., 2002), including in the early stages of the evaluation, through enhanced participative approaches. The difference observed in France may therefore be attributed to the Neo-Weberian nature of the French state (Dunn & Miller, 2007; Lynn, 2008; Pollitt & Bouckaert, 2011), where the public institutions' attempt to preserve their traditional apparatus hinders the development of more organic management styles required to foster truly participative processes. It remains that early-stage participation and its impact on project's ownership appears a key differentiation between SIM as advocated by AVISE and PE as carried out by the SGMAP.

3.2.2. The question of the outcomes

Outcomes are a central question in both PE and SIM literature. Thus most contributions in both disciplines highlight the difference between outputs and outcomes (Fouquet, 2014; Mulgan, 2010; J. Nicholls et al., 2009; Olsen & Galimidi, 2008; Rossi et al., 2003). In France, while some of the documentation proposed by AVISE does not tackle clearly the difference between outputs and outcomes (e.g. AVISE et al., 2013; CSESS, 2011), most of it stresses the distinction (e.g. European Commission / GECES, 2014; EVPA, 2015; Stievenart & Pache, 2014). AVISE, like most SIM actors, thereby emphasizes the importance of outcomes in the evaluation process and highlight the potential pitfall of mistaking outputs for outcomes.

This approach is however absent of the French PE side. A good illustration of this observation is that the "public action value chain" (cf.

Figure 2) proposed by the SGMAP goes directly from “actions” to “outcomes”²¹, without proposing any equivalent to outputs. What the methodological framing of the SGMAP suggests is also confirmed by the actual evaluation practices in the context of the MAP: the engagement letters of the five selected PE cases never hint at the difference between outcomes and outputs while one of the three evaluation reports available (vocational training for unemployed individuals) touches upon it, suggesting in its synthesis that benefits of the intervention have to be considered on a broader scale than just the simple re-integration in the labour force.

In fact, outcomes do not appear central in the MAP PEs: extracting the main evaluation questions in the five engagement letters, we get to a total of 21 questions, each focusing on one or more performance criteria as framed by the SGMAP in the “public action value chain”. Eight of these questions (38%) focus, to a certain extent, on efficiency or efficacy, the two performance criteria related to outcomes in the SGAMP’s framework (e.g. “how does the reality of inclusion for incarcerated individuals relate to the means invested in it”). On another hand, 13 of these questions (62%) focus, to a certain extent, on coherence issues (whether internal or external) and have a particular interest in processes (e.g. “how are activities structured at different territorial scales to promote inclusion”).

More broadly speaking, the term “impact” is only specifically mentioned once in the 21 identified questions. This does not mean that PEs carried out in the MAP are not interested in outcomes. They simply appear more keen on understanding the processes (in order to improve them) and the use of

²¹ “Résultats” in French, a term equated to outcomes in all of AVISE documentation making a difference between outcomes and outputs (which are then called “produits”, “réalisations” or “performance”).

resources (in order to make them more efficient) leading to the outcomes rather than proving or testing the robustness of these outcomes. This, we argue, is a key gap between PE and SIM, where a considerable part of the SIM effort is to actually verify, document, and often try to quantify the value of the outcomes.

3.2.3 Another perspective on metrics and indicators

SIM is typically carried out by relatively small organisations with limited resources. This makes spending time and manpower on undertakings not related to their core activity a substantial opportunity cost. A lot of the methods for SIMs presented by AVISE have consequently been designed with these facts in mind. Although some like SROI²² may still seem complex and time consuming to an organisation willing to carry out SIM for the first time, a lot of work is being done by the creators of these methods and organisations like AVISE in France to produce pedagogical material and provide ad'hoc support to stakeholders in order to smoothen their SIM experience. Part of this effort includes designing simple mechanisms and indicators, relatively easy to fathom for a non-initiated audience (e.g. the Outcomes Star²³). In SROI's case, involving all stakeholders in the early stages of the SIM process is also a way to include them in the development of these indicators and especially on the valuation process of the outcomes.

PE is rather far from these considerations, with a wide diversity of activities to cover, often scattered on a national scale, but also with substantial resources. As exposed above, evaluation questions are multiple and seldom explicitly focused on outcomes. If metrics are used to frame an evaluation, they

²² Social Return On Investment, a method using financial proxies to value outcomes, supported by the SROI Network.

²³ A Likert scale-based system in which stakeholder can rate pre-defined indicators to gauge a programme's outcomes

will most likely be in pure economic terms, evaluating the costs associated to a given scenario (e.g. the Vocational training for unemployed individuals case), not considering avoided costs in most cases. In some situations, part of the PE's scope is to define indicators to gauge the programme's outcomes. This is the case for the Partnership between the national education institutions and the business sector evaluation, where the creation of indicators ends up being a recommendation of the actual evaluation report...

The focus on streamlined, approachable methods with accessible and rather clearly defined indicators and metrics is therefore another key differentiation characterising SIM against PE. PE is more prone to complexity, with a strong focus on processes not requiring an analytical scaffold framed by clear indicators. Furthermore, outcomes in French PEs are most often touched upon through efficiency as a performance criterion, which may lead to a focus on cost as the main metric.

4. Conclusion: how is social impact measurement challenging programme evaluation

While sharing the same theoretical and conceptual foundations as well as some features like the participative nature of the EP as highlighted by the SGMAP, SIM and PE remain very different in their execution in France. As highlighted above, the similarities between the two can easily be explained by the way SIM took inspiration from the emergence of PE in a legislative context during the late 1960s. A striking feature of this process however, and a thought we want to emphasise in this conclusion, is how SIM, as a PE practice taken by private actors from the public space is looping back into the political agenda in today's zeitgeist, as a singular example of isomorphism.

Beyond the conceptual similarities, our analysis identifies three main gaps which are (i) how the outcomes are treated in PE and SIM; (ii) how is stakeholders' participation managed and how it affects the ownership of the evaluation process by the concerned parties; and (iii) how metrics and indicators are approached. In the French context, we hypothesise that some of these devices may be due to the Neo-Weberian public management framework, where administrations are attached to their traditional apparatus and are reluctant to introduce management mechanisms moving away from the customary top-down approach. We suggest that exploring how to draw further on the Neo-Weberian framework's modernizing tenants (Lynn, 2008) to drift towards more organic forms of management (and therefore more organic forms of PE) would be an interesting research alley.

As previously hinted, another driver we attribute the observed dissimilarities to is the obvious difference in scale between PE and SIM. SIM is carried out on specific, well identified initiatives or entities, often at a local level with a limited number of individuals impacted by the assessed activities. PE is in turn often focused on nation-wide topics with a multitude of stakeholders and activities to consider. This naturally drives up the complexity of PE, making it hard for evaluators to follow the narrow analytical paths encouraged by numerous SIM methods.

So is SIM changing PE in France? Not really at this moment. The latest iteration of PE in France indubitably gets closer to SIM practices when it sets stakeholders participation as one of its core principles, but there is no evidence that the development of SIM has played a role in this evolution. It is however interesting to note that on the SIM practitioners' side, stakeholders tend to parallel SIM to PE (e.g. Branger, Gardin, Jany-Catrice, & Pinaud, 2014), or to suggest it as a prolongation of PE (e.g. CSESS, 2011). This last view is probably the most accurate, as SIM is unlikely to replace PE due to the factors we highlight in this conclusion.

But if SIM is not changing PE, it seems fair to say that it is challenging it: envisioning SIM as a prolongation of PE, to assess initiatives at the micro level, at the closest to their stakeholders, is already bringing a considerable amount of social innovation to the PE paradigm. It offers granularity in a process characterised by its macro scale and it empowers users and the actors involved in the delivery of the programmes. It also allows for the evaluation process to trickle down the strata of the society, making evaluation and its stakes better understood, more owned by its actors and has thereby potential for increasing the overall quality of PE. Perhaps the recent launch of social impact bonds²⁴ by the French Ministry of Economy is a sign of policy moving towards this direction.

As a final thought, and as we make the case for PE to take cues from SIM, the reverse observation can also hold true. This is for instance the case for the strong focus on processes observed both in the methodology proposed by the SGMAP (internal and external coherence as preeminent performance indicators) and in the actual PE practice in the context of the MAP (a majority of the questions framing the evaluations brings about processes). This focus on process is present in some SIM methods like the Measuring Impact Framework²⁵ but remain absent in most cases and may lead to overlooking some important mechanisms occurring in the creation of outcomes and impact.

As a closing note, we want to remind of some methodological considerations framing this paper. First, as this article is part of a broader research focusing on WISEs, the PE we selected to build our

²⁴ Contrats à impact social, where the government proposes to refund investments related to a social programme, upon success (a premium can be earned on the investment, as a function of how much impact is generated). A first project call was launched in March 2016 in France.

²⁵ Holistic and detailed framework encompassing management criteria, developed by the World Business Council for Sustainable Development

analysis were specifically sampled with WISE-like activities in mind. Broadening the focus of the analysis may consequently yield somewhat different outcomes. Second, we recall that while we are able to have a very concrete understanding of how PE is carried out in France in practice due to the transparency standards pushed by the MAP, the same cannot be said for SIM, where cases data is not publically available, except for handpicked examples chosen by the stakeholders promoting SIM (hence exposing us to biased selectivity). The gap analysis may consequently be perceived as unbalanced due the absence of a real comparison between theory and practice for SIM in France. A forthcoming field research will help mitigate this limitation and test the set of hypotheses developed in this paper.

- Albareda, L., Lozano, J. M., & Ysa, T. (2007). Public policies on corporate social responsibility: The role of governments in Europe. *Journal of Business Ethics*, 74(4), 391–407.
- Assemblée Nationale. (2016). Constitution of October 4, 1958. Retrieved from <http://www2.assemblee-nationale.fr/langues/welcome-to-the-english-website-of-the-french-national-assembly>
- Austin, J., Stevenson, H., & Wei-Skillern, J. (2006). Social and Commercial Entrepreneurship: Same, Different, or Both? *Entrepreneurship Theory and Practice*, 30(1), 1–22.
- AVISE, ESSEC, & Mouves. (2013). *Petit précis de l'évaluation de l'impact social*. AVISE.
- Battesti, J.-P., Bondaz, M., Marigeaud, M., & Destais, N. (2012, December). Cadrage méthodologique de l'évaluation des politiques publiques partenariales. Inspection Générale de l'Administration, Inspection Générale des Finances, Inspection Générale des Affaires Sociales. Retrieved from <http://www.ladocumentationfrancaise.fr/var/storage/rapports-publics/124000683.pdf>
- Borus, M. E. (1979). Measuring the Impact of Employment-Related Social Programs. A Primer on the Evaluation of Employment and Training, Vocational Education, Vocational Rehabilitation, and Other Job-Oriented Programs. Retrieved from <http://eric.ed.gov/?id=ED173586>
- Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal*, 9(2), 27–40.
- Branger, V., Gardin, L., Jany-Catrice, F., & Pinaud, S. (2014). *Evaluer l'impact social d'une entreprise sociale : points de repère*. Corus' ESS.
- Broadbent, J., & Laughlin, R. (2003). Public private partnerships: an introduction. *Accounting, Auditing & Accountability Journal*, 16(3), 332–341.
- Burdge, R. J., & Vanclay, F. (1995). Social Impact Assessment: a contribution to the state of the art series. *Environmental and Social Impact Assessment*, 14.
- Clark, C., Rosenzweig, W., Long, D., & Olsen, S. (2004). Double bottom line project report: assessing social impact in double bottom line ventures.
- CSESS (Conseil Supérieur de l'ESS). (2011). *La mesure de l'impact social*. CSESS.
- Dauphin, S. (2012). Focus–Mesurer l'utilité sociale des associations. *Informations Sociales*, (4), 54–57.
- Davister, C., Defourny, J., & Grégoire, O. (2004). *Work integration social enterprises in the European Union: an overview of existing models*. Working Papers Series.
- Dees, J. G., Ewing, M., Haas, M., & Haas, P. (1998). *The meaning of social entrepreneurship*.
- Duclos, H. (2007). *Evaluer l'utilité sociale de son activité. Conduire une démarche d'auto-évaluation* (Baton Rouge / Polyprint). Paris: AVISE.
- Dunn, W. N., & Miller, D. Y. (2007). A critique of the new public management and the neo-Weberian state: advancing a critical theory of administrative reform. *Public Organization Review*, 7(4), 345–358.
- Edwards, P., & Shaoul, J. (2003). Partnerships: for better, for worse? *Accounting, Auditing & Accountability Journal*, 16(3), 397–421.
- Emerson, J. (2003). The blended value proposition. *California Management Review*, 45(4), 36.
- ESSEC. (2011). *Guide du retour social sur investissement (SROI)*.
- European Commission / GECES. (2014). *Approches proposées pour la mesure de l'impact social*. European Commission.
- European Commission, & OECD. (2013). Policy Brief on Social Entrepreneurship. Entrepreneurial activities in Europe. Publication Office of the European Union.

- Evers, A., & Laville, J.-L. (2004). *The third sector in Europe*. Edward Elgar Publishing.
- EVPA (European Venture Philanthropy Association). (2015). *Un guide pratique pour la mesure et la gestion de l'impact*. EVPA.
- Fouquet, A. (2014). L'évaluation des politiques publiques: État (s) de l'art et controverses. *Revue Française D'administration Publique*, (4), 835–847.
- Furubo, J.-E., Rist, R. C., & Sandahl, R. (2002). *International atlas of evaluation*. Transaction Publishers.
- G8's Social Impact Measurement Taskforce. (2014). *Measuring Impact*. G8.
- Gibbon, J., & Dey, C. (2011). Developments in social impact measurement in the third sector: scaling up or dumbing down? *Social and Environmental Accountability Journal*, 31(1), 63–72.
- Glaser, B., & Strauss, A. (1967). *The discovery grounded theory: strategies for qualitative inquiry*. Aldin, Chicago.
- Grieco, C. (2015). *Assessing Social Impact of Social Enterprises: Does One Size Really Fit All?* Springer. Retrieved from <http://link.springer.com/content/pdf/10.1007/978-3-319-15314-8.pdf>
- Grimsey, D., & Lewis, M. (2007). *Public private partnerships: The worldwide revolution in infrastructure provision and project finance*. Edward Elgar Publishing.
- Guskey, T. R. (1999). *Evaluating professional development*. Corwin Press.
- Heckman, J. J., LaLonde, R. J., & Smith, J. A. (1999). The economics and econometrics of active labor market programs. *Handbook of Labor Economics*, 3, 1865–2097.
- Hogan, R. L. (2010). The Historical Development of Program Evaluation: Exploring Past and Present. *Online Journal for Workforce Education and Development*, 2(4), 5.
- Hood, C. (1991). A public management for all seasons? *Public Administration*, 69(1), 3–19.
- Hulgård, L. (2011). Social economy and social enterprise: an emerging alternative to mainstream market economy? *China Journal of Social Work*, 4(3), 201–215.
- Imbens, G. W., & Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, 47(1), 5–86.
- Lapsley, I. (2009). New Public Management: The Cruellest Invention of the Human Spirit? 1. *Abacus*, 45(1), 1–21.
- Levitan, S. A., & Wurzburg, G. K. (1979). *Evaluating federal social programs: An uncertain art*. WE Upjohn Inst for.
- Lynn, L. (2008). What is a Neo-Weberian State? reflections on a Concept and its Implications. *The NISPAcee Journal of Public Administration and Policy*, 1(2), 17–30.
- Maas, K., & Liket, K. (2011). Social Impact Measurement: Classification of Methods. In R. Burritt, S. Schaltegger, M. Bennett, T. Pohjola, & M. Csutora (Eds.), *Environmental Management Accounting and Supply Chain Management* (pp. 171–202). Springer Netherlands.
- Mair, J., & Martí, I. (2006). Social entrepreneurship research: A source of explanation, prediction, and delight. *Journal of World Business*, 41(1), 36–44.
- Mulgan, G. (2010). Measuring Social Value, in" Stanford Social Innovation Review. Summer.
- Nicholls, A. (2010). The legitimacy of social entrepreneurship: reflexive isomorphism in a pre-paradigmatic field. *Entrepreneurship Theory and Practice*, 34(4), 611–633.
- Nicholls, J. (2007). Why measuring and communicating social value can help social enterprise become more competitive. London: Cabinet Office, Office of the Third Sector.
- Nicholls, J., Cupitt, S., & Durie, S. (2009). *A guide to Social Return on Investment*. Cabinet Office, Office of the Third Sector.

- Nyssens, M., & Defourny, J. (2012). *The EMES approach of social enterprise in a comparative perspective*. EMES, Working Papers.
- Olsen, S., & Galimidi, B. (2008). Catalog of Approaches to Impact Measurement - Assessing social impact in private ventures. *The Rockefeller Foundation*.
- Perret, B. (2006). De l'échec de la rationalisation des choix budgétaires (RCB) à la loi organique relative aux lois de finances (LOLF). *Revue Française D'administration Publique*, (1), 31–41.
- Pestoff, V. A. (2008). *A democratic architecture for the welfare state*. Routledge.
- Poister, T. H., Levitan, S. A., Wurzburg, G., Abert, J. G., Sobell, L. C., Sobell, M. B., ... others. (1981). *Program Evaluation: State of the Art/State of the Practice*. JSTOR.
- Pollitt, C. (1993). Occasional excursions: a brief history of policy evaluation in the UK. *Parliamentary Affairs*, 46(3), 353–363.
- Pollitt, C., & Bouckaert, G. (2011). *Public Management Reform: A comparative analysis-new public management, governance, and the Neo-Weberian state*. Oxford University Press.
- Rexel Fondation, & Improve. (2014). *Guide de la mesure d'impact social*.
- Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2003). *Evaluation: A systematic approach*. Sage publications.
- Rubin, D. B. (1973). Matching to remove bias in observational studies. *Biometrics*, 159–183.
- Scriven, M. (1991). *Evaluation thesaurus*. Sage.
- Scriven, M. (1996). The theory behind practical evaluation. *Evaluation*, 2(4), 393–404.
- SGMAP. (2015, April). Evaluation des politiques publiques (EPP): principes, processus et méthode. Retrieved from http://www.modernisation.gouv.fr/sites/default/files/epp_map_memo.pdf
- Stievenart, E., & Pache, A.-C. (2014). Evaluer l'impact social d'une entreprise sociale: points de repère. *Revue Internationale de L'économie Sociale: Recma*, (331), 76–92.
- Stufflebeam, D. (2001). Evaluation models. *New Directions for Evaluation*, 2001(89), 7–98.
- Yarbrough, D. B., Shulha, L. M., Hopson, R. K., & Caruthers, F. A. (2010). *The program evaluation standards: A guide for evaluators and evaluation users*. Sage Publications.
- Yin, R. K. (2009). *Case study research: Design and methods*. Sage publications.
- Yunus, M. (2006). Social business entrepreneurs are the solution. *Social Entrepreneurship: New Models of Sustainable Social Change*, 39–44.