Vega A, Chiasson M.

[Towards a comprehensive framework for the evaluation of small and medium enterprise policy](http://dx.doi.org/10.1177/1356389015593357).

Evaluation 2015, 21(3), 359-275.

Towards a comprehensive framework for the evaluation of small and medium enterprise policy

Arturo Vega
Business School, Newcastle University, United Kingdom

Mike Chiasson
Faculty of Management, British Columbia University, Canada

Corresponding author:
Arturo Vega. 5 Barrack Road, Newcastle upon Tyne, NE1 4SE, UK.
Email: arturo.vega@newcastle.ac.uk

Towards a comprehensive framework for the evaluation of small and medium enterprise policy

Abstract

This research demonstrates the relevance of the evaluative cycle and its diverse methodological designs in small and medium enterprise (SME) policy. We structure our arguments based on the most common phases of the cycle, namely policy justification, needs, policy theory, implementation, impact and efficiency assessments. We use an in-depth case study of public assistance to an SME to illustrate how findings from these phases go beyond the results of the additionality practice in SME policy. We employ the findings as starting points to discuss several methodological designs for the evaluation of entire programmes, policies and systems.

Introduction

SME policy has become an important element of most governments in attempting to create national competitive advantage (e.g. Audretsch et al., 2007; Shapira, 2010). There have been ambitious initiatives to support SMEs in diverse areas such as technology diffusion, internationalization and entrepreneurship. Nevertheless, all the stages of the SME policy process have received numerous critics (e.g. Dannreuther, 1999; Johnson 2005; Vega et al., 2013). This tension makes more vital the function of evaluation in the SME policy arena (e.g. Bryson et al., 1999; Curran, 2000; Edwards et al., 2007).

Most of the research and practice of SME policy evaluation have focused on the impact of programmes (Curran and Blackburn, 2001; Curran and Storey, 2002; Lenihan et al., 2007; Rideout and Grey, 2013). Trying to expand the evaluative scope, Stame (2004, 2010) pointed out that evaluation should give more attention to 'failures' in the design and implementation of programmes in order to open the 'black box' between inputs and outputs. In a seminal contribution, Pawson and Tilley (1997: xvi) affirmed that evaluation should explain 'why a programme works, for whom and in what circumstances'. From a participatory perspective, Kuhlmann (1998) argued that evaluation should go beyond impact measurement to enable 'intelligent negotiations' among actors who pursue different interests and game strategies.

Therefore, the contrast between a comprehensive evaluative approach and the current focus of SME policy evaluation opens up major opportunities to introduce different evaluation phases including policy justification, needs, policy theory, implementation, impact and efficiency assessments (e.g. McDavid et al., 2012; Patton, 2002; Posavac, 2011; Rossi et al., 2004), as well as various methodological designs to carry out this important agenda (e.g. Pawson and Tilley, 1997; Stern et al., 2012).

The paper is structured as follows. We start by considering the gaps between the evaluation phases and the mainstream work in SME policy. We then explain our case study methodology for the empirical part of the research. After that, we present a case study of the public assistance provided to an SME and analyse it through the evaluation phases. Then, we explain the use of various methodological designs in evaluating whole programmes, policies and systems. We finish by discussing further methodological implications and defining a research roadmap.

Opportunities for SME policy

The aim of this section is to determine the research and practical possibilities in SME policy evaluation. We start by explaining the phases and dynamics of the evaluative cycle. We then focus on the approaches that have been used in the SME domain. Finally, we identify gaps and avenues for the evaluation of SME policies.

The evaluative cycle

Policies can be evaluated from multiple and related perspectives. Table 1 presents an adaptation of the most common phases of the evaluative cycle, a brief explanation of each of them and some examples of their associated questions (e.g. McDavid et al., 2012; Patton, 2002; Posavac, 2011; Rossi et al., 2004). The evaluative cycle is better understood if we explain its dynamics and connection with the policy process.

Table 1. The evaluative cycle

Evaluation Phase	Explanation	Examples of Questions
Policy Justification Evaluation	Questions the mere existence of the policy	• Is the social issue a problem or an opportunity? • Is the social issue a priority for government and other stakeholders? • Has government the ability to intervene the social issue?
Needs Evaluation	Review of the social issue that is subject of the intervention	• What are the nature and structure of the social issue? • What are the gaps that need to be reduced or eliminated? • What is the distribution of the social issue?
Policy Theory Evaluation	Critical analysis of the policy design	• Are the policy components addressing the needs? • Is the interaction of different policy instruments necessary? • Are there other policy alternatives to address the needs?
Implementation Evaluation	Verification of the quality of the implementation	• Is the delivery according to the policy theory? • Are the policy workers taking the right decisions in the delivery? • Is the policy correctly adapted to the local context?
Impact Evaluation	Extent that the social issue was improved	• Have the companies accomplished their business initiatives? • Are there unintended outcomes? • What components of the policy were successful?

Efficiency Evaluation	Financial analysis	- Is the policy cost-beneficial? - Are there better policy alternatives in financial terms? - Is the policy technically efficient?

In principle, there is a logical sequence across the phases (e.g. Posavac, 2011; Rossi et al., 2004). For example, the development of any evaluation would make sense if a policy was initially justified. The policy theory should be evaluated after knowing the social needs. Similarly, the services delivered in the implementation have to be compared with the policy theory. The impact evaluation would be meaningful when the previous phases had acceptable results. Finally, the efficiency assessment has to consider the impact evaluation and the implementation costs.

The application of the phases also depends on the policy process in which the evaluations are embedded. The policy process embraces the stages of agenda-setting, problem definition, policy design, implementation, evaluation and termination (e.g. Karlsson and Conner, 2006). Evaluation is pervasive because it is not restricted to a particular timeframe but could be present in the entire process, and from different perspectives (e.g. Jann and Wegrich, 2006; Pancer and Westhues, 1989).

For instance, an ex-ante evaluation could be done on a pilot programme to support the policy process stages of agenda-setting, problem definition and policy design. This independent assessment would be based on real interventions and embrace the policy justification, needs and policy theory evaluations. There could also be a mid-term

evaluation when the programme is deployed at large-scale in order to review how it is implemented according to local contexts. Finally, an ex-post evaluation could be employed when the implementation is well-advanced. This could be done with the impact and efficiency assessments, which could lead to the termination of the programme.

Furthermore, the evaluative cycle could disrupt the policy process itself because 'evaluation can lead to diverse patterns of policy learning with different implications in terms of feedback mechanisms and a potential restart of the policy process' (Jann and Wegrich, 2006: 54). As evaluation affects the policy process and the policy process affects evaluation, each evaluation phase could directly or indirectly prompt the other phases, even in changing directions and without a fixed starting point. This makes the evaluative cycle nonlinear and complex as an open system (McDavid et al., 2012).

For example, the implementation evaluation could detect that the programme services require high levels of specialisation, which the programme workers could not offer. As a response, policy-makers could change the policy by choosing one of the options previously formulated during the policy design in order to match needs and resources. This would require the evaluation of a new implementation. The alternative option could have little impact, which would demand new evaluations of needs and policy theory. Finally, the results of these assessments could trigger new policy designs and restart the process or, conversely, undermine the justification of the policy.

To conclude, it is possible to omit or partially complete one or more evaluation phases if there were aspects of the policy or programme which were already known or could be assumed with confidence (e.g. Rossi et al., 2004). However, the inappropriate selection of evaluation phases would produce misleading results (e.g. Wholey, 2004), which stresses the relevance of reviewing the current focus of SME policy evaluation on impact measurement.

Research and practice in SME policy evaluation

The aim of the impact evaluation is to determine the extent to which the policies generate the intended improvements in society (e.g. White, 2010). The type of impact, or outcome, should be measurable and aligned with the policy being assessed (Harrison and Leitch, 1996). For instance, consultancy services to SMEs for the adoption of an innovation could be evaluated based on the effect of the programme on the number of SMEs that correctly implement and use the innovation. However, SME policy evaluation has mostly focused on the safeguarding and increase in sales and jobs (Hart, 2007; Lenihan, 2011), which is called additionality.

Storey (2000) summarised the approaches used to monitor programmes and quantify additionality in SME policy. The first two levels measure the take-up and satisfaction with the services. Level three is the opinion of the owner-managers about the additionality generated by the programmes. The next levels include control groups to compare results with counterfactual scenarios. Level four incorporates typical firms as

control groups. Level five considers SMEs that match the programme participants in observable aspects such as sector, size and location. Level six uses sophisticated statistical techniques to take out the effect of selection bias. Selection bias happens when the SMEs could have succeeded without public support.

There have been some plausible efforts to use the higher levels to measure additionality in SME policy (c.f. Lenihan, 2011). The calculation of the net effect in society could be more comprehensive by including displacement, leakage, learning spillovers, unintended outcomes, multipliers and double-counting (c.f. Chrisman and McMullan, 2002; Munday and Williams, 2009).

There are very few contemporary studies that have tried to diversify the evaluation of SME policies. Certainly, most of them have used qualitative methods. For instance, Patton et al. (2003) addressed implementation issues by explaining how a national policy on support for business planning was adapted to local needs, delivery capacities and antecedents from past experiences.

Krupnik (2012) focused on the factors affecting different actors involved in the implementation of a programme to subsidise investments. The factors -the fear of returning the funds to the European Union, programme targets and the lack of political power of the SME community- were crucial for the decision-taking, the minor effects of the subsidies and ultimately the continuation of the programme.

Finally, Beckinsale and Ram (2006) studied a programme for the adoption of information systems by ethnic minorities. They emphasised the necessity of more comprehensive services to overcome adoption barriers. Given the diversity within the sector, they determined a group of elements to improve the implementation including pre-existing relationships, cultural awareness and brokerage expertise.

Gaps and opportunities

The first gap is that most of the questions of the cycle and its dynamics have been practically untouched by policy-makers and the SME literature. This generates a series of opportunities in relation to the contemporary roles of evaluation. To begin, the phases are in line with the formative purpose of evaluation (Scriven, 1991). For instance, the inclusion of the needs, policy theory and implementation evaluations is intended to explain causal links in order to improve policy mechanisms and ultimately their impact (e.g. Dahler-Larsen, 2005, Stern et al., 2012).

The evaluative cycle can also be instrumental in driving social interactions and change during the policy process (e.g. Greene, 2006). For example, the themes and flows of the cycle would enrich the negotiation between parties who have different interests and perspectives (e.g. Kuhlmann, 1998, Patton, 1996), which is a critical need in SME policy (e.g. Johnson, 2005; Massa and Testa, 2008). Also, much of the required information tends to be qualitative and come from the SMEs themselves in the form of

objectives, needs and priorities, so guaranteeing their ample participation in evaluation (e.g. Cousins and Earl, 1992).

Furthermore, we believe that the formative, negotiating and participative character of the cycle would align the work of all the policy participants towards the goals of the SMEs, which would strongly contribute to the empowerment of the SME representation (e.g. Fetterman, 1994). This is vital since they have been consistently disregarded in the policy process (e.g. Dannreuther, 1999; Moran, 2009).

The second gap is a consequence of the first. The use of the phases implies the employment of different methodologies (e.g. Pawson and Tilley, 1997; Stern et al., 2012). The paper illustrates this diversity by presenting a thorough case-based evaluation at the assistance level and using the findings to explain different methodological designs for the evaluation of entire programmes, policies and systems. As the paper is an initial attempt to broaden the scope of SME policy evaluation, the whole exercise focuses on presenting the results and methodological discussion for each phase as well as important links between them, but does not enter into the intricate complexity of the dynamics and timing of the cycle.

Research design

As mentioned, the methodology employed is a case study. We examined the programme assistance provided to an SME for the adoption of an information system (IS). A case

study is defined by Robson (2002:178) as 'an empirical investigation of a particular phenomenon within its real life context using multiple sources of evidence'. Yin (2014:16) adds that 'the boundaries between phenomenon and context may not be clearly evident'. Accordingly, a feature of this research is that there were several units of analysis, i.e. phenomena under investigation, and diverse groups and contexts which were interconnected in several ways. This is known as an embedded case design because the units and contexts must be linked to develop conclusions about the assistance and the programme (e.g. Scholz and Tietje, 2002).

For example, we found that the policy justification was affected by problems in the demand for the services, which were detected in the implementation evaluation. The unit of analysis in the needs evaluation was the initiative of the SME, and the context embraced some SME characteristics such as the knowledge of the decision-maker and financial resources. In the policy theory evaluation, the unit of analysis was the programme design, which was compared with the needs of the SME. The behaviour of the programme workers (PWs) was the unit of analysis in the implementation evaluation, while the needs of the SME and some programme characteristics such as targets and resources were the context. The impact evaluation was made on the basis of the result of the initiative of the SME together with an unintended outcome. Lastly, the efficiency evaluation was affected by some programme practices which were identified during the implementation evaluation.

There is a clear aspect that makes vital the use of several sources and data collection methods in evaluation so as to guarantee validity, namely the sensitivity of the information to the participants. In fact, many SMEs depend on public assistance to accomplish their business initiatives, which may bias their views in favour of programmes (e.g. Bryson et al., 1999). It could also be the other way around because SMEs could undervalue the impact of public support to demand further assistance (e.g. Curran, 2000). Regarding PWs, they may confuse programme evaluation with personal evaluation. They may also think that the programme could be terminated or that personal information could be misused. All these concerns can induce defensive behaviours (Posavac, 2011).

Therefore, we employed diverse methods and sources to triangulate information (e.g. Patton, 2002), including interviews with the SME personnel and PWs, and review of numerous material such as economic policy documents, manuals for the management of policies and public funds, policy and funding reports, the programme organisation proposal to access public funds, programme reports and programme assistance files. We also observed the use of the IS in the SME, read internet information about the SME, the programme organisation and the IS, and had informal conversations with the participants.

We also guaranteed the validity of the information with the research process itself and through the dynamics of the interviews. In the interaction with the SME personnel we

emphasised that the research was to understand SME innovation and improve programme services rather than to evaluate programmes. The conversations were related to the SME initiative and not to programme interventions. The issues related to the intervention spontaneously emerged during the interviews.

It is important to mention that the initial reaction of the SME personnel was to overestimate the value of the programme support. However, once the researchers took the control of the interviews as explained above, the accounts became more realistic. For example, after a positive description of the programme work, an SME informant expressed the view that the system developed by the programme was unfinished and the SME could not use it. We triangulated this information by seeing the SME employees working with the old system, verifying the existence of bugs in the new development as well as a lack of data in its master files.

The aspects that positively influenced our interaction with the PWs were that they knew we had conscientiously prepared ourselves for the interviews and had already talked in detail with the SME personnel. These tactics encouraged PWs to be forthright about the pros and cons of their programmes and contexts. They explained deficiencies that were related to their own work, such as their predisposition to focus on numerical targets and not on qualitative aspects. This kind of 'mea culpa' substantially strengthens the validity of the case study.

The case information

This section is the background information and chronological narrative (Yin, 2014) of the assistance provided by the programme organisation ICTASSIST in the adoption of an IS by the SME LanguagesCo, which was done through the public programme PP-ICTServe. We use pseudonyms to identify the participants.

We begin by explaining some characteristics of the programme organisation, for example its portfolio of services and human resources. We then describe some issues presented in the implementation of its programmes, such as the priorities given to certain SMEs and the relations with other programme organisations. After that, we describe the programme used in the case in terms of its services, target SMEs and so on. We continue by explaining the SME needs and giving an account of the initiative. Finally, we report the events and outcomes that occurred during the assistance such as the interaction between the programme and SME personnel and the deliverables.

Programme organisation

ICTASSIST is a unit belonging to the departments of computing and communications at a university. The programme organisation has been running public programmes for SMEs since its creation in 2002. The services delivered included collaborative research and development for information and communication technology (ICT) companies, the location of ICT companies in the ICTASSIST premises, ICT advice, IS design and development, and ICT and IS training. Each programme of ICTASSIST delivered a specific subset of these services.

The human resource structure of ICTASSIST was composed of a programme manager, project officers for the administrative activities, academics and students to deliver the services, and third-party providers for the carrying out of some standard services.

Characteristics in the implementation

Many of the programmes had a delayed start, often caused by the administrative procedures of the funding bodies. In turn, ICTASSIST had delayed the recruitment of a number of its personnel until the signing of the contracts, which had left a reduced time to accomplish the targets. Effectively, according to the manuals for the management of public funds, only the Higher Education Innovation Fund made it compulsory to start the programmes before the signing of the contracts. To handle this, ICTASSIST had to employ extra project officers and sub-contract more third-party services to catch up once the programmes were up and running.

An important consideration was that many SMEs saw the programmes as their last option for success. The programme manager suggested we take advantage of this to get interviews with the SMEs' personnel. A report of ICTASSIST illustrated this dependency:

'Many of the SMEs had developed their ICT capabilities in-house without any dedicated ICT expertise. Their systems were often fragmented and unstable… A suspicion of the ICT sector resulted in a number of companies experiencing

bottlenecks… In addition, the majority of the companies simply did not have the financial resources or time to invest in ICT.'

The range of time given per assistance was broad, varying from 2 weeks to 4 months, including administrative work and any third-party services sub-contracted by the programmes. For example, the design and development of a webpage by an external provider could take 2 weeks, while the delivery of a student summer project could take 4 months. We verified this by reading the programme assistance files. The programme manager commented:

'[The scale of service] is mostly based on what the company wants but also on how this business will affect the programme outputs… The more outputs we can generate in terms of business growth, which are our targets under the ERDF [European Regional Development Fund], the more assistance we give.'

The ICTASSIST personnel appeared to use a simplistic method to select and service SMEs, often changing the scope of action of the programmes to match the demand and attain programme targets, for example by changing the services and the people in charge of the delivery. The programme manager and a project officer shared similar opinions on these aspects. The programme manager commented:

'The pressure to catch up with the targets meant that we were just trying to find as many companies as we could, not to say "no" to anybody because we were so desperate to get companies signed up to assist.'

'For example, one of the programmes focused on [collaborative research and development with ICT SMEs]… [But] it is a really difficult task to match an SME with an academic, in terms of something the academics will give to the SMEs as products [to commercialise]… [In addition,] none of the academics were interested in any of these [standard] works [required by the non-ICT SMEs] because they all considered it to be low level, so I was told by the head of the department, "If you want [standard] assistance choose student projects".'

Finally, the programme manager observed that after the ICTASSIST interventions, the role of the programmes was just to recommend to the SMEs what should be done from an ICT perspective. The programme personnel completely distanced themselves from business-related issues. But even for ICT subjects, the ICTASSIST personnel did not specify third-party providers in order to avoid responsibility for the quality of any additional work. Moreover, the programme manager evaded responsibility regarding post-service problems or further SME needs. A project officer of the unit illustrated his passive approach to both linking SMEs with other providers and following-up on the SME initiatives in this way:

'Sometimes an SME can be interested in some marketing work with the management school, or sometimes in an engineering project... [We refer the client to another programme of the University] if they mention a problem and a programme that could be suitable for them.'

'I wouldn't chase the company and follow-up [the SME initiative] because it means more work, ha, ha… I didn't do any follow-up, not really.'

Programme PP-ICTServe

This programme was designed to deliver high level knowledge exchange between academics in the departments of computing and communications and SMEs from the ICT sector. The idea was to help these SMEs to develop high-tech products to commercialise to other companies. PP-ICTServe was 50% funded by the ERDF with matched funding from the university.

The ERDF required the quantity of SMEs assisted and the increase and safeguarding of sales and jobs for evaluation purposes. Around six months after the programme interventions, the SME personnel had to fill out a form indicating the increased and safeguarded jobs and sales generated since the end of the assistance. In cases where the business initiative of an SME was assisted by more than one programme, the contribution to the targets had to be allocated proportionally to the number of programmes.

SME and the IS initiative

LanguagesCo is a one-stop language shop for different types of services, such as translation, interpretation and subtitling by linking service providers and clients. The company had eight employees, with the production director the only person with some IS expertise including a basic knowledge of database programming and some implementation experience with an enterprise system.

The company had an application developed in Access by the production director himself. The system had evolved over time according to business requirements, but suffered improvised modifications. As a result, the application had inconsistent information across database tables, and the speed was dropping as different types of operations were added. Additionally, there were key features whose development was outstanding, including the extension of the supplier database and the integration of some functions. Thus, the initiative of the SME was to develop a new system.

LanguagesCo decided to start its project in April 2004. The production director said that the assistance of ICTassist was crucial to accomplish the company initiative, given his basic knowledge and limited time to personally undertake the project.

The application was delivered in October 2004. However, as of the last quarter of 2006, pending tasks included the correction of some functionality and the migration of the historical data from the old system to the new one, which effectively made the new system inoperable. In a radical change of plans, the production director decided to

modify the old system based on improvements in the new system, but he only managed to finish a few tasks. He explained:

> 'I'm bringing into the old system parts taken from the new prototype… I am taking elements of the design and the concepts, so for example going from the single supplier table to the modular supplier tables.'

The production director thought that with additional public support he would have preferred to have debugged the new system and completed a data migration.

Assistance process and deliverables

The general opinion of the production director regarding the service of the programme was good:

> 'The guy they put on was very, very good, I mean, he produced a good working prototype with a lot of very useful stuff in it.'

However, if we analyse the complete interview and other material, there is evidence of mixed value in the assistance. To begin, PP-ICTServe initially used an academic supervisor and a student to serve LanguagesCo through a summer project. Nevertheless, the director stated that over a period of three months the progress of the university's personnel was not as agreed between the parties. He also disagreed with a recommendation of using PHP as the development platform because no one in the SME

had expertise with this scripting language, and it would have created a dependence on third-party providers for the maintenance and extensions to the system. The production director pointed out:

> 'The first part, with the student, took too long… My brief there, which was in May, that's when the person of the University started to work… I basically fell out with them I think in August time. That's the point when I realised that they couldn't do what I wanted them to do… It was annoying.'

As a result, the client and the programme manager of ICTASSIST changed the direction of the assistance. They agreed to redevelop the system in Access and have the hands-on work completed by a project officer from PP-ICTServe. We confirmed this in the interviews with the programme personnel and from reading the programme assistance files. The project officer was an expert practitioner of IS, with a first degree in computing and professional background in enterprise systems. The new system was finally developed and tested. As stated, some corrections to the code and the migration of the historical data were never completed. In fact, we observed the bugs, an empty database and the employees using the old system.

The programme manager and the project officer appeared to be distant from the client's situation after the intervention. For example, the project officer answered the following regarding the connection to other service providers and follow-up activities:

'The idea was that the production director was to take it further, he would carry on the work and modify it… It was running ok, it was a while ago, I haven't been in touch with [LanguagesCo] since then.'

To finish, the ERDF evaluation form was not filled out by the client because the system was not in operation. However, the production director wrote a letter explaining his expectation of increasing sales by 20% as soon as the application was up and running.

Analysis and evaluation of results

This section is an initial attempt to address the evaluative cycle. The exercise will focus on the first two questions of each phase in table 1. The results will be used in the next section as starting points to discuss more inclusive methodological designs for the evaluation of policies, programmes and systems.

Policy justification evaluation

The programme had clear problems in meeting its objective. The fact that the programme personnel refocused the design during the implementation is a strong indicator that the university support to develop innovative products with ICT SMEs *was probably not an opportunity*. A potential explanation for this is that the activities in the geographical area of the university have been dominated by farming and the university business itself. The region is not a technological cluster such as Cambridge or Sussex.

Additionally, both the academics and the personnel of the ICT SMEs were not willing to work with each other because each had *different and conflicting priorities*. Academics wanted to create knowledge to enhance their careers through leading edge research and the SMEs were interested in developing products to make profit often requiring conventional services.

Needs evaluation

LanguagesCo needed to accomplish *several tasks to adopt the IS*. Basically, the SME had to select products, services and suppliers, had to plan and manage a project, to analyse, design and develop the system, to test and debug it, to migrate the data to the new system, prepare technical documentation, train users and technical people, put the system into operation, and carry out the maintenance and extensions to the system.

However, there were *important gaps* between the adoption tasks and the internal capabilities of the SME. The company did not have the skills and resources to select services and suppliers, to design and develop the system, to test and debug it, to migrate data, to prepare the technical documentation, to train technical people and put the system into operation. The SME could have taken greater charge of the maintenance and extensions to the system after proper technical training.

Policy theory evaluation

The focus of PP-ICTServe was to deliver high level knowledge exchange between the academics of the departments of computing and communications and ICT SMEs in order to develop innovative products to commercialise to other companies. We can appreciate that the *programme components* of PP-ICTServe were not originally designed to address the needs of LanguagesCo. First, LanguagesCo required conventional IS services, and not high level knowledge exchange. Second, these services required technicians with expertise in mainstream IS, and not academics from the departments of computing or communications. Third, LanguagesCo is an intermediary of language services, and not an ICT SME. Fourth, the system to manage the interaction between language service providers and clients was for internal use, and not to commercialise to other companies.

Finally, the programme personnel should be able to *interact with other programmes or private suppliers* when the programme does not cover some or all the SME needs, as this case required. In other words, any limitation in what a programme can offer to clients should be addressed by other programmes or the capabilities in the market.

Implementation evaluation

PWs should select SMEs, design specific services for them, deliver the services, connect the SMEs with other providers, and follow-up on the SME initiatives for which assistance was provided. We evaluated the implementation by analysing each of these activities to assess the appropriateness of *what the PWs did or did not do and the*

reasons for this. As part of this analysis it is essential to determine *if the services delivered were modified from the original programme design.*

To begin, PP-ICTServe could have rejected LanguagesCo as a recipient of public services because there was little correspondence between the needs of the SME and the programme aims. However, the SME was selected primarily due to the potential contribution of the company to the programme targets. At the same time, the problems with the targets were generated by a mismatch between the intended programme offer and the market demand as well as the delayed start of the programmes given the tardy signature of contracts between the policy administrators and ICTASSIST.

Given the complications with the high-level knowledge exchange services, the programme manager offered mainstream IS services, which were undertaken by students under academic supervision. As a result, the services for LanguagesCo included project planning and management, system analysis, design and development, tests and debugging. Nevertheless, the services were incomplete because they did not include other SME needs such as data migration, technical documentation, training and putting the system into operation.

After a slow and problematic delivery, the parties decided to remove the student and the supervisor from the project and give responsibility to a project officer for the development of the hands-on activities. Despite the good work, this person ran out of time and could not complete the tests and debugging. So, the system was never finished.

Alternatively, the programme personnel could have connected the SME with other programmes to address its other needs, but this was not done. A plausible explanation is that the programme personnel wanted to avoid the apportioning of outputs with other programmes. Finally, there was no attempt to follow-up on the SME initiative. We believe that both connection and follow-up would have been on the agenda if the final success of the SME initiatives was considered in the formal evaluations. However, the programme was focused on the increase and safeguarding of jobs and sales instead.

Impact evaluation

In this part of the evaluation, we do not intend to make a numerical analysis of the impact generated by the programme intervention, but to give some insights that could help to understand and improve it. First of all, the SME *did not finish the development of the IS and consequently is not using it*. This is a compelling reason to drastically reduce the purported additionality. We also detected a relevant *unintended outcome*, namely the extra and ineffective attempts of the production director to modify the old system from the design of the unfinished development of the programme. The time invested in this decision had a negative effect on the impact generated by the assistance.

Efficiency evaluation

As with the impact evaluation, we do not intend to determine a value for the efficiency of the programme but to identify factors that should be included in the calculations.

With this aim, we found that the PWs delivered something that did not match the primary purpose of the programme. Arguably, the decision of a programme manager should not replace the decisions taken by policy-making partnerships, which are aided by advisors, economic research and documented past experiences. On the other hand, experienced PWs could be knowledgeable in the needs of the SMEs and could adapt a programme to balance needs, resources and work pressures. In any case, the opportunity cost associated with the modifications of the policies by the implementers should be incorporated in the analysis. It is a different type of counterfactual scenario which compares the *financial ratios of the intended and realised programmes.*

Towards inclusive methodological designs

In this section we progress from the evaluation of a public intervention to the evaluation of a whole programme, policy and system, with the aim of explaining methodological diversity in SME policy evaluation. We will exemplify potential approaches to address the first two questions of each evaluation phase in table 1.

For the policy justification, our findings indicate that the programme had serious problems in meeting its targets. This information could be complemented with secondary data such as the historical regional variation in the number of ICT SMEs and ICT patents. In addition, the numbers could be compared with other jurisdictions to provide a broader contrast. The objective would be to determine the relationship between different sets of information to strengthen the results (e.g. Greene 2007;

Tarrow, 2004), which is known as the convergent parallel approach (Creswell, 2014). If the region is doing well, probably *the social issue is not an opportunity* (Edquist, 2002). If the region is lagging behind, the evaluative cycle could continue depending on the answer to other questions in the policy justification evaluation, for example if *the social issue is a priority for relevant stakeholders* (e.g. Karlsson and Conner, 2006).

Regarding the needs evaluation, it could start by developing a number of case studies with some of the ICT SMEs that were assisted or approached by the programme so as to analyse the needs in the target companies. The cases should be differentiated by region, SME size and the type of initiative in order to understand common and dissimilar needs. The case studies could be developed using a similar approach to the one presented in this paper. This could be relevant as guidance to the design of a survey instrument to do *a general verification of the needs* (e.g. Tarrow, 2004). Creswell (2014) calls this approach the exploratory sequential design because the qualitative data helps to identify important issues which are confirmed or refined with the use of wide-reaching quantitative methods.

This part of the needs evaluation could indicate, for example, that one of the needs in the SMEs is marketing expertise to commercialise new products. However, this is an observable need, which could have many potential underlying causes. It is the underlying causes that have to be addressed by the policies, for instance the deficiency

of university courses on marketing, the migration of professionals to other regions or the unwillingness of SMEs to trust external advice.

We can appreciate that the needs evaluation should represent a multi-level research process *to detect systemic failures within the structure of a system of innovation* (e.g. Arnold, 2004; Georghiou, 1998; Kuhlmann, 1998). For this reason, these studies require the use of methodological pluralism and multiple disciplines as the evaluation progresses and new findings, relations and directions emerge (e.g. Vega and Brown, 2011). This technique is based on a deep ontological perspective and resembles the retroductive approach used in realist evaluations (Pawson and Tilley, 1997).

A characteristic of the policy theory evaluation is that there could be *many policy alternatives to address each cause* detected in the needs evaluation (Vega et al., 2008). For instance, if an underlying cause is that professionals migrate to other regions, policy-makers could decide to improve the quality of life in the region, to soften the immigration laws for professionals, or link university expertise in marketing to SMEs. Some measures directly address the issue of migration, for example the improvement in quality of life, whereas others indirectly counteract its effects, such as the link of university expertise to SMEs.

A further complexity in the policy theory evaluation is that *each policy instrument could be designed in many forms*, and not all of them would be effective for SMEs. For instance, the link of university expertise and SMEs could focus on market research,

product launch or international marketing. The type of services could include consultancy, advice or coaching. Universities could also sub-contract third-party providers or jointly work with other public programmes. In addition, the interaction with the SMEs could be via traditional face-to-face methods or electronic platforms.

If the policy was a support programme, case studies like ours would be a good beginning because they scrutinise the match between the needs of different SMEs and the programme theory components, as well as the effectiveness (Yin, 2014). For example, one of our findings was that the programme services did not match or matched inadequately the SME's needs, which implied the need for joint work with other public or private providers.

This reinforces the systemic view of SME policy from a different angle in the sense that the evaluation of a system implies *the evaluation of more than one policy mechanism -* for instance programmes, regulations, the development of technological infrastructure and the re-engineering of public administrative processes- in which each should have its own evaluative criteria but also all must be oriented towards a common goal (e.g. Arnold, 2004; Klijn, 2005; Pawson, 2007; Rogers, 2008). Therefore, further formative research would be necessary to evaluate a working policy system (e.g. Nutley et al., 2003). Certainly, this type of evaluation would embrace a research process and mixed methods.

Multi-level complexity also influences the implementation evaluation. We found that PWs used *considerable discretion to change the programme scope* as well as took some *biased decisions* primarily to accomplish the targets. These findings are not uncommon in several public service areas. Actually, there is substantial research on these topics (e.g. Lipsky, 2010; Maynard-Moody and Musheno, 2003), including studies with a focus on enterprise policy (c.f. Vega et al., 2008, 2013). The underlying causes of discretion and biased PWs are similar and quite persistent (e.g. high targets, low resources, conflict of interests and the evaluations themselves) and from a pragmatic position, we believe, once these behaviours are detected it should not be the focus of evaluators, but more the responsibility of policy-makers to take action, if possible.

The impact evaluation could start with a quantitative study by experimenting with control groups and using statistical techniques to reduce the selection bias. Nevertheless, the value of the qualitative methods is that they bring life to the numbers by explaining their form (e.g. Posavac, 2011; Tarrow, 2004). For example, we found that *the SME did not adopt the IS* and *was also affected by unintended outcomes*. This would certainly explain a negative impact. Accordingly, this approach is called the explanatory sequential design (Creswell, 2014). The case studies should be discriminated by different attributes such as sector, SME size, types of initiatives, or the most and least successful participants to better interpret the results. As SME policy has

to be systemic, the analysis should include all the policies which affected the adoption of IS in the SMEs of a geographical area.

To conclude, the efficiency evaluation should rely on *proper impact and implementation evaluations to calculate the cost-benefit ratio*, including all the policy mechanisms affecting the adoption of IS by the SMEs in the region. But the efficiency evaluation is also a comparison with other policy alternatives. The policy theory is especially relevant here because it proposes a group of policy instruments to address the same underlying needs affecting the SMEs. The other alternatives should be used as *benchmarks to define if the programme was the best option.*

However, we have to take into account that the programme was reinterpreted by the PWs and there is a mismatch between the intended and realised delivery. In situations like this, the financial evaluation should also consider the opportunity cost of the service actually delivered versus the service as initially designed.

Conclusions

This study has produced two relevant contributions, namely the demonstration of the value of the evaluative cycle in SME policy and the use of diverse methodological designs to carry out this comprehensive work. To begin, we validated that the evaluation phases answer crucial questions that different stakeholders must know in

order to guide the dynamics of complex and messy negotiations (e.g. Kuhlmann, 1998, Patton, 1996).

As appreciated in the case study and the methodological discussion, the SMEs have to actively participate throughout the cycle (e.g. Cousins and Earl, 1992) as sources of concepts that shape the whole evaluation. Furthermore, the alignment of all the phases with the success of the SMEs and the assessment of each policy component to this end close potential doors to circumvent the true focus of evaluation, which would empower the SME sector in the policy process (e.g. Fetterman, 1994).

Although the inclusion of a pervasive evaluation cycle and its associated questions is a necessary requirement to improve negotiations, participation and empowerment in SME policy, this is insufficient to guarantee effective change. The phases must also be undertaken properly in order to be really formative (Scriven, 1991), which stresses the importance of the methodological designs.

We support the view that formative evaluations must be approached in a systemic way (e.g. Arnold, 2004; Georghiou, 1998; Kuhlmann, 1998). As explained in the paper, the systemic nature of SME policy can be depicted from many angles. Firstly, a group of SMEs could have several common needs to be more competitive. Secondly, each need could have various underlying causes. Thirdly, each cause could be addressed by more than one policy instrument. Fourthly, different policy instruments could directly remove

a cause or ameliorate its effects at different points in the system. Fifthly, a policy instrument could be implemented by more than one organisation.

We also exemplified diverse mixed-methods to undertake each evaluation phase, such as the convergent parallel approach to strengthen results, the explorative sequential design to generalise from fuzzy social issues, and the explanatory sequential design to understand the composition of widespread indicators (e.g. Creswell, 2014; Greene, 2007; Tarrow, 2004). However, the systemic nature of SME policy means that these and other combinations of methods would have to be used several times according to the emerging needs of multi-level evaluations.

A further consequence of the evaluation of systems is a move from individual programmes to policy systems as main units of analysis (Stern, 2006; Stern et al., 2012). It does not make sense to evaluate a programme in isolation but instead a group of policy mechanisms which together should contribute to the same objective (e.g. Arnold, 2004; Klijn, 2005; Pawson, 2007; Rogers, 2008). Moreover, the systems perspective supports the view that evaluation must focus on the explanation of the non-linear contribution of a group of policy mechanisms instead of the attribution of additionality to a single instrument (Mayne, 2001; Stern et al., 2012).

The policy system should initially be decomposed into different sub-systems (Arnold, 1994) according to the diverse needs of the SMEs and the corresponding ramifications in terms of underlying causes, policy mechanisms and implementers. This would

require extensive methodological pluralism to detect issues at any point in the system. Nevertheless, evaluators must not lose the focus on the main goal (e.g. Stern et al., 2012). Therefore, the sub-systems must be recomposed to validate the contribution of each instrument to the causal chain of the main policy theory.

This study opens up a plethora of opportunities for SME policy evaluation. However, the practice of the evaluative cycle is challenging because it is mostly unknown and complex, apart from being considerably expensive (e.g. Bamberger et al., 2006). We propose a research agenda to ameliorate these constraints, including systemic goals and attribution, the methodologies for each phase, their associated questions, the cycle dynamics, the omission or partial development of the phases, participation and empowerment in the cycle, the cycle as a learning and negotiation tool, and the diffusion of the cycle to different stakeholders. We believe that this research is a firm step in this direction.

Funding

References

Arnold E (2004) Evaluating research and innovation policy: A systems world needs systems evaluations. *Research Evaluation* 13(1): 3-17.

Audretsch D, Grilo I and Thurik R (2007) Explaining entrepreneurship and the role of policy: A framework. In: Audretsch D, Grilo I and Thurik A (eds) *Handbook of Research on Entrepreneurship Policy*. Cheltenham: Edward Elgar Publisher, pp. 1-17.

Bamberger M, Rugh J and Marby L (2006) *Real World Evaluation: Working Under Budget, Time, Data and Political Constraints*. California: Sage

Beckinsale M and Ram M (2006) Delivering ICT to ethnic minority businesses: An action-research approach. *Environment and Planning C: Government and Policy* 24(6): 847-867.

Bryson J, Daniels P and Ingram D (1999) Evaluating the impact of Business Links on the performance and profitability of SMEs in the United Kingdom. *Policy Studies* 20(2): 95-105.

Creswell J (2014) *Research Design: Qualitative, Quantitative and Mixed Methods Approaches*, 4rd edn. California: Sage.

Chrisman J and McMullan W (2002) Some additional comments on the sources and measurements of benefits of small business assistance programs. *Journal of Small Business Management* 40(1): 43-50.

Cousins J and Earl M (1992) The case for participatory evaluation. *Educational Evaluation and Policy Analysis* 14(4): 397-418.

Curran J (2000) What is small business policy in the UK for? Evaluation and assessing small business policies. *International Small Business Journal* 18(3): 36-50.

Curran J and Blackburn R (2001) *Researching the Small Enterprise*. London: Sage.

Curran J and Storey D (2002) Small business policy in the United Kingdom: The inheritance of the small business service and implications for its future effectiveness. *Environment and Planning C: Government and Policy* 20(2): 163-177.

Dahler-Larsen P (2005) Evaluation and public management. In: Ferlie E, Lynn L and Pollitt C (eds) *The Oxford Handbook of Public Management*. Oxford: Oxford University press, pp. 615-639.

Dannreuther C (1999) Discrete dialogues and the legitimation of EU SME policy. *Journal of European Public Policy* 6(3): 436-455.

Edquist C (2002) Innovation policy - A systemic approach. In: Archibugi D and Lundvall B (eds) *The Globalizing Learning Economy*. Oxford: Oxford University Press, pp. 219-238.

Edwards T, Delbridge R and Munday M (2007) A critical assessment of the evaluation of EU interventions for innovation in the SME sector in Wales. *Urban Studies* 44(12): 2429-2447.

Fetterman D (1994) Empowerment evaluation. *Evaluation Practice* 15(1): 1-15.

Georghiou L (1998) Issues in the evaluation of innovation and technology policy. *Evaluation* 4(1): 37-51.

Greene J (2006) Evaluation, democracy and social change. In: Shaw I, Greene J and Mark M (eds) *The SAGE Handbook of Evaluation*. London: Sage, pp. 118-140.

Greene J (1997) Participatory evaluation. In: Mabry L (ed) *Evaluation and the Post-Modern Dilemma - Advances in program Evaluation (vol. 3)*. Greenwich: JAI Press, pp. 171-189.

Greene J (2007) *Mixed Methods in Social Inquiry*. San Francisco: Josey-Bass.

Harrison T and Leitch C (1996) Whatever you hit call the target: An alternative approach to small business. In: Danson M (ed) *Small Firm Foundation and Regional Economic Development*. London: Routledge, pp. 221-247.

Hart M (2007) Evaluating EU regional policy: How might we understand the causal connections between interventions and outcomes more effectively? *Policy Studies* 28(4): 295-308.

Jann W and Wegrich K (2006) Theories of the policy cycle. In: Fischer F, Miller G and Sidney M (eds) *Handbook of Public Policy Analysis: Theory, Politics and Methods*. Florida: CRC Press, pp. 43-62.

Johnson S (2005) SME support policy: Efficiency, equity, ideology or vote-seeking? In: *Institute of Small Business and Entrepreneurship (ISBE) conference*, Blackpool, UK, 1-3 November.

Karlsson O and Conner R (2006) The relationship between evaluation and politics. In: Shaw I, Greene J and Mark M (eds) *The SAGE Handbook of Evaluation*. London: Sage, pp. 225-242.

Klijn E (2005) Networks and inter-organizational management: Challenging, steering, evaluation, and the role of public actors in public management. In: Ferlie E, Lynn L and Pollitt C (eds) *The Oxford Handbook of Public Management*. Oxford: Oxford University press, pp. 257-281.

Krupnik S (2012) Evaluating the social mechanism of investment subsidies using an abductive approach. *Evaluation* 18(4): 466-476.

Kuhlmann S (1998) Moderation of policy-making?: Science and technology policy evaluation beyond impact measurement – The case of Germany. *Evaluation* 4(2): 130-148.

Lenihan H (2011) Enterprise policy evaluation: Is there a 'new' way of doing it? *Evaluation and Program Planning* 34(4): 232-332.

Lenihan H, Hart M and Roper S (2007) Industrial policy evaluation: Theoretical foundations and empirical innovations: New wine in new bottles. *International Review of Applied Economics* 21(3): 313-319.

Lipsky M (2010) *Street-Level Bureaucracy: The Dilemmas of Individuals in Public Service*, 2nd edn. New York: Russell Sage Foundation.

Massa S and Testa S (2008) Innovation and SMEs: Misaligned perspectives and goals among entrepreneurs, academics and policy makers. *Technovation* 28(7): 393-407.

Maynard-Moody S and Musheno M (2003) *Cops, Teachers, Counsellors: Stories from the Front Lines of Public Service*. Ann Arbor: University of Michigan Press.

Mayne J (2001) Addressing attribution through contribution analysis: Using performance measures sensibly. *Canadian Journal of Program Evaluation* 16(1), 1-24.

McDavid J Huse I and Hawthorn L (2012), *Program Evaluation and Performance Measurement: An Introduction to Practice*. California: Sage.

Moran M (2009) *Business, Politics and Society – An Anglo American Comparison*. Oxford: Oxford University Press.

Munday M and Williams G (2009) Comparing gross and net employment outcomes from structural fund programmes: Issues from the mid-term evaluations of the objective 1 programme in West Wales and the Valleys. *Evaluation* 15(2): 225-246.

Nutley S, Walter I and Davies H (2003) From knowing to doing: A framework for understanding the evidence-into-practice agenda. *Evaluation* 9(2): 125-148.

Pancer M and Westhues A (1989) A developmental stage approach to program planning and evaluation. *Evaluation Review* 13(1): 56-77.

Patton D, Marlow S and Ram M (2003) Interpretive analysis as an evaluative tool: The case of Mustard.UK.Com, a high-growth small business programme. *Environment and Planning C: Government and Policy* 21(6): 813-824.

Patton M (1981) *Practical Evaluation*. California: Sage.

Patton M (1996) *A world larger than formative and summative*. Evaluation Practice 17(2): 131-144.

Patton M (2002) *Qualitative Research and Evaluation Methods*, 3rd edn. California: Sage.

Pawson R (2007) Causality for Beginners. NCRM Research Methods Festival, Leeds University. URL: http://eprints.ncrm.ac.uk/245/

Pawson R and Tilley N (1997) *Realistic Evaluation*. London: Sage.

Posavac E (2011) *Program Evaluation*, 8th edn. New Jersey: Pearson Education.

Rideout E and Gray D (2013) Does entrepreneurship education really work? A review and methodological critique of the empirical literature on the effects of university-based entrepreneurship education. *Journal of Small Business Management* 51(3): 329-351.

Robson C (2002) *Real World Research*, 2nd edn. Oxford: Blackwell.

Rogers P (2008) Using program theories to evaluate complicated and complex aspects of interventions. *Evaluation* 14(1): 29-48.

Rossi P, Lipsey M and Freeman H (2004) *Evaluation: A Systematic Approach*, 7th edn. California: Sage.

Scholz R and Tietje O (2002) *Embedded Case Study Methods: Integrating Quantitative and Qualitative Knowledge*. California: Sage.

Scriven M (1991) Beyond formative and summative evaluation. In: McLaughlin M and Phillips D (eds) *Evaluation and Education: At Quarter Century*. Illinois: The University of Chicago Press, pp. 18-64.

Shapira P (2010) Innovation and small and midsize enterprise: Innovation dynamics and policy strategies. In: Smits R, Kuhlmann S and Shapira P (eds) *The Theory and Practice of Innovation Policy: An International Research Handbook*. Cheltenham: Edward Elgar, chapter 8.

Stame N (2004) Theory-based evaluation and types of complexity. *Evaluation* 10(1): 58-76.

Stame N (2010) What doesn't work? Three failures, many answers. *Evaluation* 16(4): 371-387.

Stern E (2006) Contextual challenges for evaluation practice. In: Shaw I, Greene J and Mark M (eds) *The SAGE Handbook of Evaluation*. London: Sage, pp. 292-314.

Stern E, Stame N, Mayne J, Forss K, Davies R and Befani B (2012) Broadening the range of designs and methods for impact evaluations. Department for International Development, working paper 38, London.

Storey D (2000) Six steps to heaven: Evaluating the impact of public policies to support small businesses in developed economies. In: Sexton D and Landstom H (eds) *The Blackwell Handbook of Entrepreneurship*. Oxford: Blackwell, pp. 176-193.

Storey D (2003) Entrepreneurship, small and medium sized enterprises and public policy. In: Acs Z and Audretsch D (eds) *Handbook of Entrepreneurship Research*. Dordrecht: Kluwer, pp. 473-511.

Tarrow S (2004) Bridging the quantitative-qualitative divide. In: Brady H and Collier D (eds.) *Rethinking Social Inquiry: diverse Tools, Shared Standards*. Maryland: Rowman and Littlefield, pp. 171-179.

Vega A and Brown D (2011) Systems of innovation, multidisciplinarity and methodological pluralism: A realist approach to guide the future of information systems research and practice. In: Chiasson M, Henfridsson O, Karsten H and DeGross J (eds) *Researching the Future in Information Systems*. New York: Springer, pp. 249-268.

Vega A, Chiasson M and Brown D (2008) Extending the research agenda on diffusion: The case of public program interventions for the adoption of e-business systems in SMEs. *Journal of Information Technology* 23(2): 109-117.

Vega A, Chiasson M and Brown D (2013) Understanding the causes of informal and formal discretion in the delivery of enterprise policies: a multiple case study. *Environment and Planning C: Government and Policy* 31(1): 102-118.

White H (2010) A contribution to current debates in impact evaluation. *Evaluation* 16(2): 153-164.

Wholey J (2004) Evaluability assessment. In: Wholey J, Hatru H and Newcomer K (eds) *Handbook of Practical Evaluation*. California: Jossey Bass, pp. 33-62.

Yin R (2014) *Case Study Research: Design and Method*, 5th edn. London: Sage.

ARTURO VEGA is a Lecturer in Innovation and Public Policy. His research is trans-disciplinary and methodologically pluralist. He worked in multinational corporations for 10 years.

MIKE CHIASSON is a Professor of Information Systems. His area of expertise is enterprise systems applied to the health and SME sectors, including policy.