Evaluating the performance and impact of intercultural policies

This paper intends to provide cities with practical tools for evaluating the performance and impact of intercultural policies. It outlines a framework for evaluating intercultural policies and discusses the issues that are crucial when designing and carrying out the evaluations.

The approach is pragmatic. While duly taking into account existing academic literature, the focus is on laying out concrete and synthetic guidelines that can help cities to design and carry out evaluation exercises in practice.

To this end, the paper is accompanied by four examples of policy evaluation setting out concrete steps and indicators for each case.

The paper is based on preparatory discussion within the network and on the Workshop held in Pécs on 20-21 June 2013.

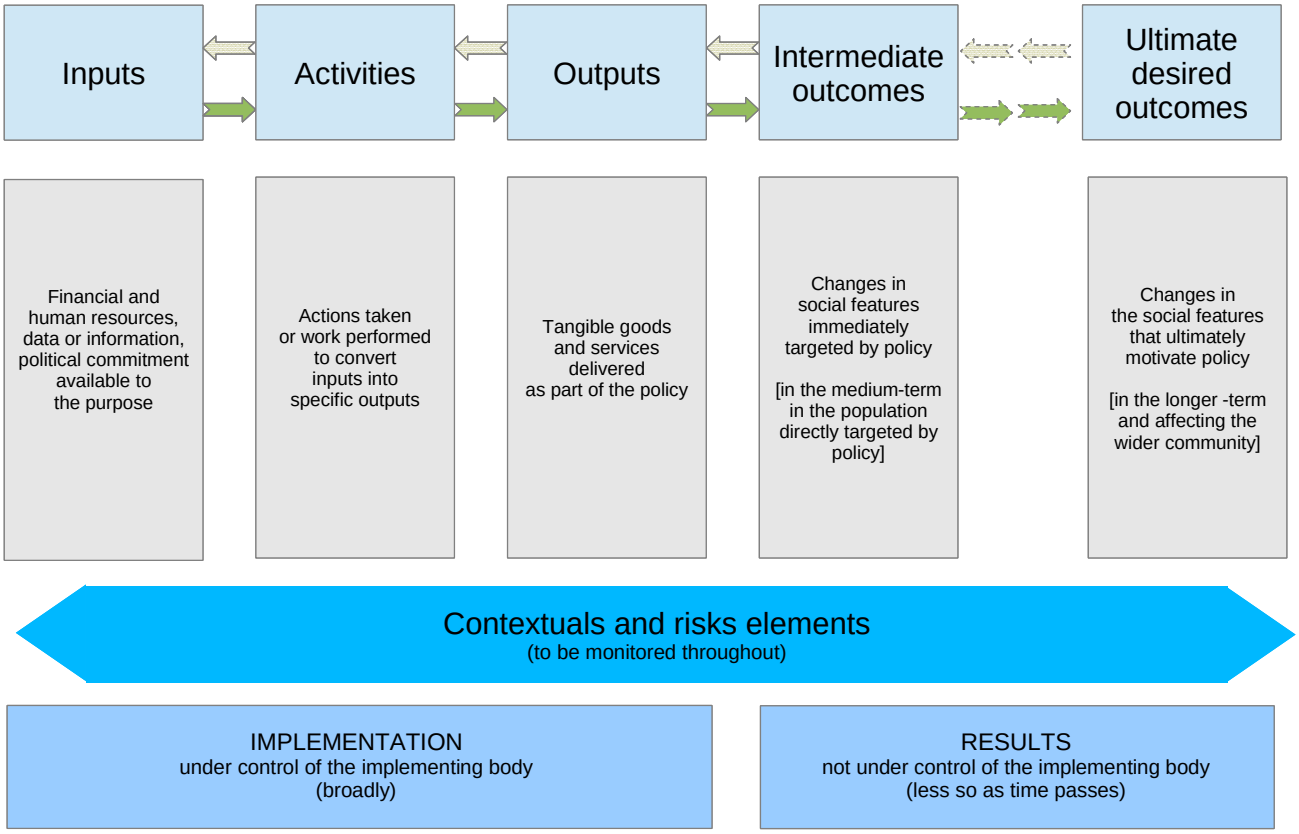
A framework for evaluation: Theory of change and results chain

Intercultural policies are set in place to deliver some type of social change. A "theory of change" is a description of how the policy is supposed to deliver the desired change.

Theory of change can be described as a road map for getting from situation A to a situation B, specifying actions, actors, resources, time lines and indicators for success. It refers to the interconnections between a vision, goals, strategies and outcomes along with the underlying theory or assumptions of causality (why are we expecting the interventions to deliver the expected changes).

There are different ways of expressing a theory of change. A simple and straightforward way is through a "results chain". A "results chain" maps the logical and plausible sequence of steps that goes from the inputs to the desired final outcome of policies.
The picture below illustrates a plausible results chain.

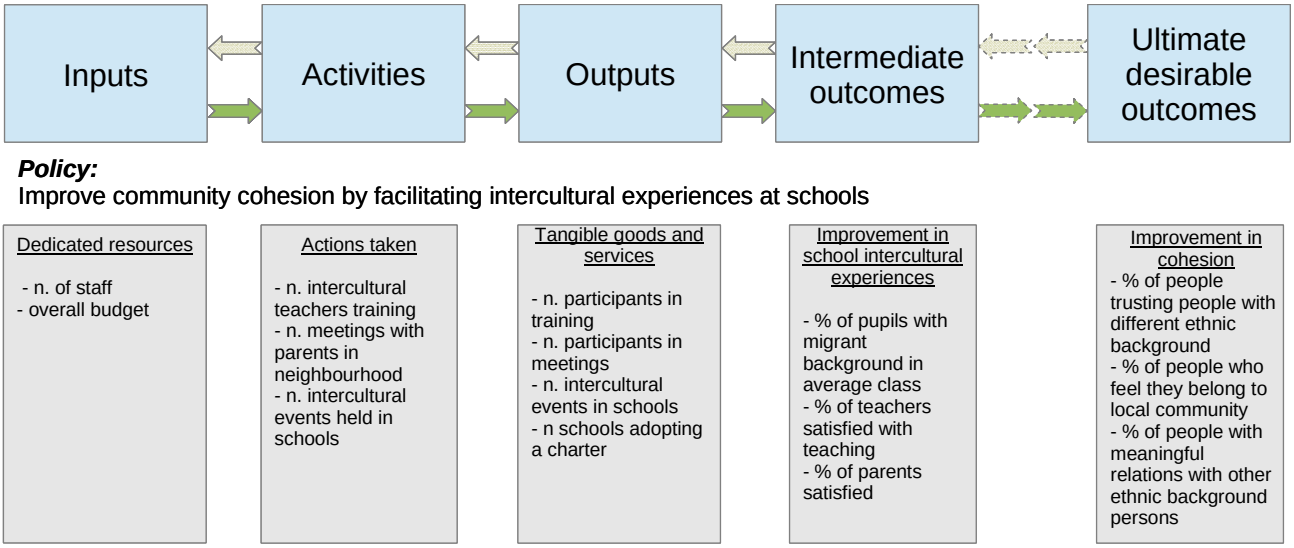
Figure 1. Theory of change: The result chain



Source(s): adapted from World Bank, 2011

It is important that, in addition to the steps of the chain, the evaluation also identifies explicitly the *contextual and risk elements,* ie, those elements of the wider context that are expected to behave in a certain way during implementation and that, if such expectations are violated, will alter the result chain and therefore the ultimate impacts of the policy (eg, if an unexpected jump in the unemployment rate takes place during a campaign to improve the perception of migrants, this will affect people's broader feelings and thus also the likely impacts of the campaign).

The table below provides a fictitious example of a policy aiming at improving cohesion by facilitating intercultural experiences at schools. The example will be used in what follows to illustrate the theoretical statements. Four concrete examples of results chain are in the Annex to this paper.



A definition of evaluation: performance and impact

The evaluation consists in measuring how outputs and outcomes change at different stages and in assessing to what extent such changes can be attributed to the policy (and its activities).

A crucial task in the evaluation is thus the identification of adequate indicators, on which data are collected at each stage. Indicators can be qualitative or quantitative. Information and data may be collected through a variety of means, including questionnaires, interviews, surveys, case records. Indicators should fulfil certain criteria: be clearly defined; be strictly related to the output/outcome to be measured; be understandable to the outside and accepted by the stakeholders; avoid inappropriate incentives.[1]

Indicators are needed both in relations to inputs and outputs (implementation) as well as intermediate and long-term outcomes (results). On the basis of this classification, two complementary definitions of evaluation exercises can be given:

1. Policy performance evaluation (ref. implementation). It measures *what is delivered (outputs) through the activities and inputs* mobilised. It can take place during and shortly after policy implementation. It is an important part of evaluation and it is the part most often carried out. However it does not say about the final impact on the variables of interest (in our example, the quality of intercultural experience and social cohesion);

[1] Cock J.C., Evaluating the Impact of Voluntary and Community Sector Organisations on Community Cohesion, Goldsmiths University of London. Downloadable at http://www.ioe.mmu.ac.uk/caec/docs/Evaluating%20the%20Impact%20of%20Voluntary%20and%20Community%20Sector%20Organisations%20on%20Community%20Cohesion.pdf.

2. Policy impact evaluation (ref. results). It measures *what changes in the (intermediate/ultimate) outcomes has the policy caused* through the outputs delivered. It requires the use of indicators that reveal the effects of policy through time. It also calls for ways of assessing causality and attribution, ie to what extent these effects are caused by the policy and not by other factors (in our example, to what extent the improvement in intercultural experiences at schools and, broaderly, in social cohesion is due to the policy and not to other contextual changes).

For most convincing results, it is necessary to assess both performance and impact, in order to understand both what works (or fails to work) and why works (or fails to work).

Integrating evaluation into policy design

The result chain is bi-directional, working respectively:

- from right to left, for the design of policies - by deriving backward from the desired ultimate outcomes (as inscribed in the strategic vision) the needed intermediate steps;
- from left to right, for the evaluation of policies, by identifying at each steps the indicators and the actors involved.

This bi-directionality requires that the evaluation is integrated from the beginning into the policy design and pursued throughout, rather than carried out ex-post. Specifically, integration will help:

- designing the policy appropriately, in particular by forcing to set out explicitly the desired policy outcomes and the intermediate steps, and thus to identify relevant assumptions at each step that would otherwise remain implicit;

- implementing the policy, as the constant monitoring will allow unexpected effects to be promptly react to and feed back into policy through regular in-between reviews that will help pinpoint specific difficulties;

- evaluating the policy, as the relevant data and information will be collected during the process and not only at the end (when data and information will be inevitably poorer);

- budgeting the policy appropriately and saving resources. It is often said that evaluation exercises are too expensive and are therefore discarded. However, the costs of evaluation may be relatively small when compared to the costs of policy, and likely to be more than counterbalanced by the savings that can be realized;

- communicating the policy, in its different steps, to the whole society

Assessing causality and attribution

Ultimate and intermediate policy outcomes are most often affected by a multiplicity of factors, which makes difficult to determine what proportion of success is caused by the policy intervention being evaluated (which is ultimately what we are interested in).

In principle, this would only be possible by comparing what actually happened with what would have happened if the policy was not in place. The latter is referred to as the "counterfactual". The counterfactual is unknown by definition and must be constructed as part of the evaluation strategy.

There are different ways of constructing a counterfactual. A common classification is as follows, from the methodologically most robust to the least:

- Randomised controlled trials. It consists in randomly assigning individual units into an intervention group (in our example: a random sample of schools which the policy is delivered to) and a control group (which the policy is not delivered to - representing the counterfactual). The effect of policy is measured by the difference in outputs and outcomes between the two groups of schools. This method is the most accurate because the random assignment ensures that the control group represents with accuracy what would have happened in the absence of policy. It is also the most costly and it is now always possible (for legal, practical or ethical reasons).
- Comparison group studies. It consists in identifying two groups with similar characteristics (in our example: the schools in two different neighborhoods) and delivering the policy only to one of the groups (in our example, the schools in only one of the neighborhoods). As before, the effect of policy is measured by the difference in outputs and outcomes between the two groups. Well-designed comparison group studies may be nearly as accurate as randomised trials. There is however always the risk of unobservable variables that affect differently the two groups (in our example, it could be the crisis of a large company located in one of the neighboorhood) and therefore bias the results.
- Single group pre-post studies. It consists in comparing outputs and outcomes before and after the policy is delivered. It assumes that no change to outputs/outcomes would have intervened in the absence of policy. This assumption is quite strong because of the many variables that could affect social outcomes during a relatively long period of time (in our example, it could be a deep recession hitting the city and inducing a deterioriation of social cohesion, which cannot be attributed to the policy – which may have rather waived the negative impact of the crisis).

The choice among the three options inevitably involves a trade-off between the costs and the opportunities of the different methodologies. If randomised trials are not feasible, comparison group studies and single group pre-post studies may be adequate

to the extent that the analysis of results takes into duly account of the inherent limitations[2].

Possible shortcuts

The ultimate desired outcomes are to a specific policy may be particularly difficult because of the many effects interacting over the relatively long period of time necessary for the policy effect to unfold fully. In that case, the evaluation exercise could be limited to assess the impact of policy on intermediate outcomes, relying on existing literature to proof the impact on long-term outcomes.

In our examples above, assessing the impact of the policy on cohesion would be difficult, as cohesion is a very complex concept depending on many factors.

However, suppose that some studies have established that improving the quality of intercultural experience at schools also increases social cohesion.
Then, evaluation could be limited to demonstrate the (positive) impact of the policy on intercultural experiences at schools, and this will be sufficient to conclude that it also has a (positive) impact on cohesion.[3]

Other issues in policy evaluation

1. Preparatory work

Impact evaluation is expensive as it requires the collection of significant amounts of data over longer periods of time, and sometimes sophisticated statistical analysis. For this reason it is worth evaluating the impact only of policies that mobilise important resources, deal with strategic issues, are innovative, untested and likely to affect policy-making.
Before undertaking impact evaluation, at least the following steps should be undertaken:

- a through review of existing literature on similar cases, which would help structuring and carrying out the evaluation and interpreting its results (and in some cases may advise to abandon the policy itself)
- a through review of data and information available, also in relation to the counterfactual.
- to plan the evaluation prospectively, ahead of implementation, including frequency of data collection, the sources, the way to do it (who will do it), and the required resources.

---

[2] More detailed discussion is in:
World Bank (2011), Impact evaluation in practice, downloadable at www.worldbank.org/ieinpractice.
European Commission (2011), A guide to social experimentation, downloadable at ec.europa.eu/social/BlobServlet?docId=7102&langId=en;
Americal Evaluation Association (2004) What Constitutes Strong Evidence of a Program's Effectiveness?, downloadable at www.whitehouse.gov/omb/part/2004_program_eval.pdf.

[3] Cock J.C. *cit*.

2. Qualitative information and stakeholder views

The focus of the paper is on indicators (primarily quantitative). However qualitative information, monitoring data, and process evaluations are also needed for program implementation and to examine questions of process that are critical to informing and interpreting the results from impact evaluations.

Furthermore, it is crucial that stakeholders views are considered from the very start (when setting out the theory of change, the results chain and indicators) to the end (the analysis of results). Stakeholder views could also give important insights during the process of evaluation.

3. Unintended effects

The policy could have side-effects (positive or negative) on variables that are not targeted. It is important to design the evaluation process in a way as to be able to capture such unintended positive or negative effects of the intervention. This may require thinking in advance to additional indicators that are not immediately related to our (long-term or intermediate) objectives. It could be a way of including stakeholders preoccupations or negative views into the design of the evaluation.

4. Considering the costs of policy

Performance and impact evaluation assess what the policy has achieved (in terms, respectively, of outputs or outcomes), ie its effectiveness in achieving the desired outcomes.

When comparing alternative strategies, it could be important to bring the costs into the picture. The concepts of cost-effectiveness and efficiency becomes relevant:

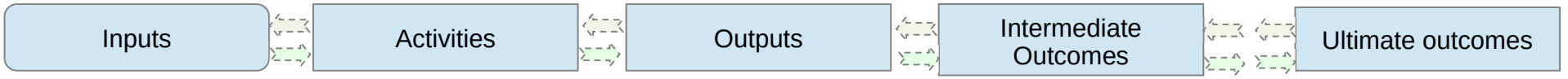
- cost effective strategy: delivers the desired outcome at least cost.
- efficient strategy: maximises the benefits to costs ratio. This requires monetising the changes in outcomes (benefits) and inputs (costs) enacted by policy (with cost-benefit analysis techniques). Cost-benefit analysis is widely applied in environmental policy. It may become difficult in social/cultural policy.

Annex: 4 real-life examples

We provide here two examples of policies to be evaluated and the relative results chain (with indicators to be used for performance/impact evaluation).

Evaluating own goals and processes (Copenhagen)

Copenhagen is extending the use of 'theory of change' to the evaluation of own groups goals and processes (after using for assessing external projects since 2007). This initiative is based on the understanding that policy will create change only if ownership is already established in the development phase. The exercise has a cross-sectoral mandate and involves multiple stakeholders in a process of co-creation. The evaluation of processes and of policies proceed intertwined to the common objectives of securing ownership, ensuring adequate feedbacks into the policy loop and strengthening the knowledge base.



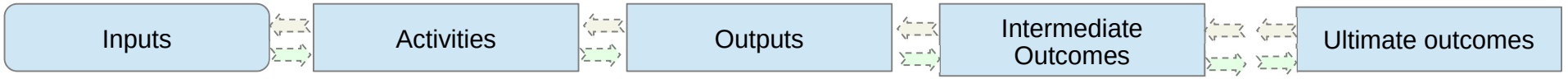
Describe/quantify the resources available	Describe/quantify the activities performed	Describe/quantify the outputs of the activities	Describe/quantify to intermediate outcomes	Describe/quantify to ultimate outcomes
[*human, financial informational and political resources]*	*[activities to involve civil society and other admins and to strengthen knowledge base]*	*[success in involving admins and civil society and improvement of knowledge base]*	*[success in developing a knowledge-solid and shared policy]*	*[social inclusion and cohesion in Copenhagen]*
- Personnel and offices in the department in charge, in the mayor office and in other administrations [n. full-time equivalent, office space and equipment] - Dedicate budget [€, by dpt/admin if appropriate] - Civil society organisations and other external partners (research institutions, ministries, cities) to be involved [qualitative description] - Political commitment [qualitative description]	- Involve all city dpt/admins [n. meetings/contacts] - Engage with civil society organisations and public [n. workshop/confs; n. persons involved; other] - Monitor city inclusion efforts [qualitative description; key insights] - Facilitate expert Think-Thanks with policy focus [n. persons and funding involved; qualitative description of activities] - Participate in research/international networks [n. workshs/confs;. participants involved; n. presentations given]	- Networks and action plans [network groups with other admin/levels; action plans/status report from all admin; outputs from bilateral meetings – qualitative descriptions] - Civil society boards [number; partners involved; n. people directly and indirectly reached; outputs (feedback,comments)] - Expert advices [description of Think Thank advice – qualitative] - Knowledge gained [examples from other cities; other insights - qualitative]	- Policy is developed in time [qualitative description] - Ownership is felt across city administrations, civil society organisation and public at large [% persons satisfied, in city admin, civil society organisation, city at large - survey-based] - The policy is built on a strong research base [feedback from the research community - qualitative description]	- Trust [% of people trusting members of other groups, etc - survey-based] - Discrimination [reported cases of discrimination, etc] - Equal opportunities [wage gaps unemployment differentials in labour markets; differential outcomes at school; etc] - Sense of belonging [%persons feeling at home in a diverse city, etc – survey based] - Any other dimension of cohesion/inclusion deemed relevant

Context and risk factors

Changes in the social, economic, political, context that may affect the impact of policy
(eg. changes in policy priorities and budgetary allocations; negative/positive economic shocks)

Innovation and growth in an Intercultural strategy (Patras)

Patras is currently developing its intercultural strategy. The strategy is multidimensional, encompassing the economic, cultural and political domains. The overall aim is to develop Patras as an intercultural meeting point building on the geographical position and history of the city. Under the economic domain (the focus of this chart) the Strategy emphasises the potential of migration to boost the economic and innovation potential of the city. Activities to facilitate migrants' entrepreneurship and participation to the labour markets are foreseen.



Describe/quantify the resources available

[human, financial informational and political resources]

- Personnel [n. full-time equivalent] and Budget [€]
- Dedicated offices and equipment [location, n. PC/other equipment]
- Volunteers available to contribute [number, charateristics]
- Potential and engagement of local NGOs [number, qualitative description of interest]
- Engagement of local businesses associations [qualitative description]
- Political commitment [qualitative description]
- Engagement of Patras "Council for the Integration of Migrants" [number of sessions, qualitative description of interest

Describe/quantify the activities performed

[activities to engage, meet, train and promote local and migrant entrepreneurs/workers]

- Engage migrants and other residents/actors [n. meetings with local and migrants associations, NGOs, university and science parks]
- Promote encountering [n. workshops, conferences and participants – also in origin countries]
- Provide training [n. courses for migrant entrepreneurs and workers, languages, culture, health, etc.]
- Promoting excellence [n. prizes and applicants]
- Other activities [qualitative/quantitative description]

Describe/quantify the outputs of the activities

[success with local and migrants entrepreneurs and workers and recognition in society]

- Engaging [n. clauses in formal docs; n. bilateral agreements with NGOs, univ/science parks; local migrants associations]
- Encountering [n. participants directly/ indirectly reached, incl. origin countries; n. formal relationships emerged]
- Training [satisfaction of participants and local employers, survey-based]
- Promoting excellence [recognition by migrants and broader society, survey-based]
- Recognition of the strategy in public debate [presence in news etc]
- Engagement of related local bodies e.g. Chamber of Commerce, Regional Admionistration Unit of Western Greece

Describe/quantify to intermediate outcomes

[contribution of migrant entrepreneurs and workers to the local economy]

- Migrant entrepreuners [n. firms, by sector, and related employment]
- their contribution to innovation [% of innovative firms; % in science parks]
- their contribution to exports [n. migrants' firms exporting abroad and % to total exports]
- Contribution of migrants to local economy [% in local employment, by skills and occupation; % of executive, management and research position, % migrants in local PA, by position]

Describe/quantify to ultimate outcomes

[economic development and competitiveness; social cohesion]

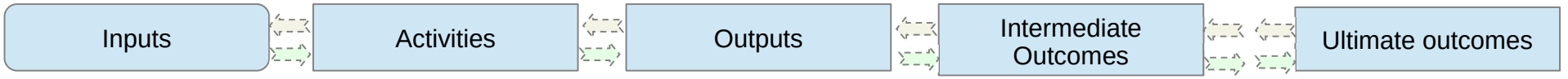
- Change in the local economy performance [growth of value added, employment, exports; technological and innovation content]
- Change in perception of role of migrants in local economy [qualitative, survey-based]
- Evolution in trust across and between groups [% of people trusting members of other groups, survey-based]

Context and risk factors

Changes in the social, economic, political, context that may affect the impact of policy, as currently being the case
(eg. changes in policy priorities and budgetary allocations, in/out-migration patterns; negative/positive economic shocks)

Welcome Centre (Hamburg)

The Welcome Centre was established in 2007 in the context of Welcome to Hamburg Programme. The core target group are students and qualified workers, to whom the Centre provides residence permit and registration services. The Centre also provides to all migrants basic information helping start their life in Hamburg (ie, how to find a house, open a bank account, etc). A web portal complements the activities of the Centre. In 2012 the Centre provided over 6,500 permits (up from ca 3,000 in 2008) and over 16,000 information services (6,000 in 2007). Weblicks were 190,000.



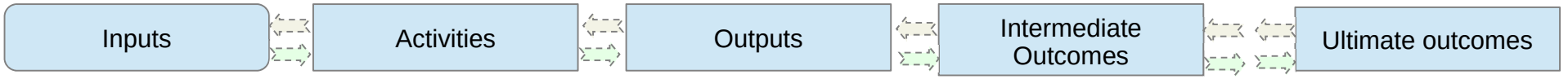
Describe/quantify the resources available [*human, financial informational and political resources]*	Describe/quantify the activities performed *[awareness, organisational and support activities; visits received]*	Describe/quantify the outputs of the activities *[services provided: residence permits, information provided]*	Describe/quantify to intermediate outcomes *[satisfaction of visitors and impact on their integration in local economic, political, cultural, civic life]*	Describe/quantify to ultimate outcomes *[intercultural integration in the city; relationship between local and migrants; trust]*
- Personnel [n. full-time equivalent] - Budget [€] - Dedicated offices and equipment [m2&location; n. PC/other equipment] - Volunteers available to contribute [n. and characteristics] - Political commitment [qualitative description]	- Awareness raising [n. leaflets, etc] - Visits received in the Centre and through Web portal [n. migrants visiting the Centre, by charactiristic (gender, nationality, education level, reason for migration); webclicks] - Portal updates and info [qualitative description] - Other events and activities during the year [qualitative description]	- Services to core group targets [n. permits, n. of registrations – by characteristics (see under 'Activities')] - Services to all migrants [n. leaflets distributed etc; qualitative description of information conveyed and to whom] - Services through Web portal [n. dowloads] - Immediate feedbacks [qualitative description: how do people have found out]	- Feedback of visitors [qualitative - by characteristics (see under 'Activities')] - Feedback of those that did not visit the Centre [survey of the reasons (eg, language, info, not useful, did not know) and consequences (feeling of exclusion?)] - Usefulness of the information [n. visitors reporting effective use] - Integration of visitors [n. employed after x months/years; n. participanting in local cultural/political life] .	- Economic integration [easiness of getting a job; unemployment in migrant population; migrants in high skill jobs; deskilling]; - Cultural and political integration [migrants in local political and cultural associations; survey of sense of belonging] - Evolution of trust between groups and openness of local population to migrants [survey based] - Decrease of discrimination [n. of reported cases]

Context and risk factors

Changes in the social, economic, political, context that may affect the impact of policy
(eg. changes in policy priorities and economic shocks)

Youth Mayor in Lewishman (London)

The Young Mayor was established 10 years ago to provide a focus and a channel for young people's views to be heard by decision makers and to ultimately increase civic participation and social cohesion. The Young Mayor has an annual budget of £30,000 and is supported by a Young Advisors' Group and a Young Citizens' Panel. The Young Mayor is between 13 and 17 years old and is democratically elected by 11-19 year olds every year through Lewisham's schools and colleges. The programme is now widely acknowledged. Most recent turnout was 52%, much higher than in the contemporary mid-term council election (30%).



Describe/quantify the resources available	Describe/quantify the activities performed	Describe/quantify the outputs of the activities	Describe/quantify to intermediate outcomes	Describe/quantify to ultimate outcomes
[*human, financial informational and political resources]*	*[awareness, organisational and support activities; training and coaching to candidates and teams]*	*[success in engaging youth; awareness and appreciation in broader population; capacity to cross gender/ethnic boundaries]*	*[youth empowerment and civic engagement; political accountability to youth; integrating 'minorities' in democratic processes]*	*[trust; local identity; social cohesion]*
- Personnel [n. full-time equivalent] - Budget [£] - Dedicated offices and equipment [m2&location, n. PC/other equipment] - Political commitment [qualitative]	- Social media [n. webpg, tweets, FB entries, blogs] - Campaigning [n. Leaflets, meetings, n. of support team members] - Teachers and other stakeholders engagement [n. meetings,…] - Training [n. hours training,…] - Meetings and events during the year [n. meetings] - Other activities [describe]	- Turnout [% of electorate] - Candidates and teams [n. and composition] - Cross-gender/ethnic group voting [% voting] - Recognition in youth and broader population [% residents aware of/appreciating the programme] - Presence in public debate [keywords in social networks; newspapers; local tv]	- Careers of former mayors/advisors [qualitative] - Turnout of younger cohort/ethnic minorities in adult elections [%, ex-post survey] - Youth influence on policy-making [n. enacted policies promoted by youth mayor/advisors; integration of youth issues in mainstream policies (qualitative)]	- Change in (perception of) role of ethnic minorities and youth in political life [qualitative, survey-based] - Evolution in trust across and between groups [% of people trusting members of other groups, survey-based] - Establishment local identity [% of people identifying with the locality rather than with the ethnic group, survey based)

Context and risk factors

Changes in the social, economic, political, context that may affect the impact of policy
(eg. In/out-migration patterns; negative/positive economic shocks; changes in political sentiment)