12. Policy Evaluation using the Framework for the Functioning of the Performing Arts in Society

This final chapter presents a model for the evaluation of municipal performing-arts policy. The model is based upon Figure 10.1, which presents the framework for describing the functioning of performing arts in society. It also incorporates the insights from the previous chapter on current cultural-policy evaluation efforts in praxis and theory. Thus, this chapter answers the last two research questions: the way in which the functioning of performing arts in society can be evaluated and measured, and the way in which municipalities can improve their cultural policy evaluation efforts.

12.1. A Model for Evaluating Performing-Arts Policy: What should be measured?

The framework for describing the functioning of the performing arts in urban society (Figure 10.1) represents three instances where the performing arts create values and functions in society: (I) the performance itself (existing within its organizational setting), which represents aesthetic values based upon the aesthetic properties of the performance and the extrinsic values that result from the existence of the cultural institutions in a city, (II) the experience of the performance by audience members, which leads to the realization of values and subsequent functions at personal level, and (III) the functioning at societal level. The previous chapter concluded that current policy evaluation practices of Dutch municipalities hardly address levels II and III, while the legitimization of cultural policy is mainly dependent on level II.

The previous chapter discussed several methods of policy evaluation that were found in the cultural policy and management literature. Although the subject of performance evaluation is far from being a mature scientific field, the discussion of the various methods did yield some useful insights. First, several relevant indicators and methods to measure them have been found. Second, it turned out to be useful to distinguish between primary objectives and secondary objectives – which is not the same as the distinction between intrinsic and extrinsic values and functions – and between values created for the professional field, audience members, and the city as a whole. These three distinctions correspond to the levels present in the model of Figure 10.1: the aesthetic values of the performing artwork itself, the
values and functions for the audience member(s) and the values and functions for society. Thus the following model for policy evaluation ensues.

The model comprises four related fields of value assessment. The experience of the aesthetic values present in the productions enables the personal values and subsequent functions. However, the professional assessment of the aesthetic value of the productions (or artistic quality) in itself is not an assessment of the values at personal level. Personal Value evaluation is a separate field within arts-policy evaluation. It involves assessing whether or not the values and functions associated with performing arts indeed have been realized for individual audience members and the audience as a whole.¹ In turn, the personal functions contribute to the realization of societal values and functions. Evaluating this chain of values and functions (from aesthetic value to personal value to societal value) comprises the primary value evaluation. But more is at stake. It is useful to evaluate the way in which this

¹ The Aesthetic Value evaluation may be regarded as assessing the possibility of aesthetic values being realized for audiences, the Personal Value evaluation involves assessment of whether or not they indeed have been realized.
12. Policy Evaluation using the Framework for the Functioning of the Performing Arts in Society

chain of value creation is organized. This involves assessing whether or not the (performing) arts institutions are able to deliver these values on a continuous basis and to a variety of audiences, preferably those audiences for whom the aesthetic expression is relevant. This comprises secondary value evaluation.

12.1.1. Aesthetic Value
The aesthetic values of the productions relate to the values mentioned in cell A of Table 10.1 (expression of ideas and perceptions and artistic development). They are relevant to cultural policy evaluation. This means that measures for tracking aesthetic value should be included in a model for the evaluation of cultural policy. As discussed in Chapter 11, such measures prevail in the current methods used by cities to evaluate their cultural policy. Measures to assess aesthetic value are also – rightfully – included in most of the theoretical approaches to cultural policy evaluation, most notably in the Public Value approach (Holden, 2004) and the professional perspective used in the threefold Balanced Scorecard as proposed by Boorsma and Chiaravalloti (2009).

The most important instrument in evaluating the cultural value is the use of independent advisory boards of experts who assess the artistic quality of the artworks produced, their position within the field of the discipline, and the contribution of the organization to the cultural system as a whole. This presupposes that cultural institutions record peer and critical reviews of their productions, and that they include data on prize money won and cooperation with other artists and arts organizations in their annual reports. The percentage of funds devoted to risky productions and the commissions of new works can be relevant measures as well. A remark must be made about the position of the advisory boards. The question posed by the city administration to the advisory board is important. The model presented here presupposes a retrospective outlook, which is not always adopted in cultural policy. This means that the advisory board evaluates past performance of an institution and uses it to formulate advice on future decisions.

12.1.2. Personal Value
The intrinsic values of the performing arts at personal level are the ‘excitement due to the experience of non-present worlds’, ‘empathizing with imagined emotions’, ‘delight in the use of the power of imagination’, ‘experience of new perceptions’ and ‘testing of one’s views and insights’. The last three result from the use of the imaginative power, i.e., the artistic nature of the experience. They are unique to aesthetic experience while the other two values can also be realized in other types of social activities. None the less, from the policy documents it follows that all five intrinsic values are relevant in the evaluation of arts policy. Currently municipalities do not assess the realization of these aesthetic values at personal level, although this is the core of what cultural policy is about. Chapter 11 suggested that the nature of the aesthetic experience and the influence of supportive services on the nature of the experience should be researched through audience surveys. For the purpose of policy evaluation, these surveys should uncover whether or not the intrinsic values of aesthetic
experience have indeed occurred for the audience members. The authors discussed in the previous chapter did not provide methods to conduct such surveys. However, within audience research, various methodologies, both qualitative and quantitative, have been developed which may shed light on the values generated. These will be discussed in Section 12.2.2.

Specifically for the performing arts, the communal nature of the experiences is relevant. At first glance, this hardly seems relevant to our research as this is an integral part of the experience of the performing arts because it is a collective activity by nature. However, audience research should include questions regarding the *experienced* communality. Do the attendees feel they have had a communal experience? Did they talk about the performance with other audience members? And does this make them feel they belong to a community (either the community that consists of the audience members actually present or the community that is represented in the aesthetic properties of the performance)?

The intrinsic *functions* on personal level (the sublimation of needs and the satisfaction of sublimated needs, and the possible change of views and insights as a result of the testing of views and insights) are related to the personal growth of individuals. Unfortunately, from a policy point of view, it is hard to imagine an evaluation method. This is because it cannot be predicted whether these effects will occur and, if they do, whether or not they are the result of an aesthetic experience of the individual. It is far more likely that they will occur as a result of many other experiences of the individual as well. So the realization of non-artistic and artistic aesthetic values seems a far more suitable measure than the occurrence of intrinsic functions. This means that, in the Personal Value field, only the potency for the realization of intrinsic functions can be evaluated. Thus, cultural policy evaluation has the following reasoning: when intrinsic values are realized for audience members – which can be assessed by means of audience research – the policy theory suggests that subsequent functions will follow at some point in time.\(^2\) However, this is not researched for policy evaluation purposes. Although the genesis of such personal functions is possible in theory researching (through longitudinal research which tracks the personal development of audience members and assesses the value of aesthetic experiences for this development through in-depth interviews) practical reasons ensure that such research is not conducted for policy evaluation.\(^3\)

The extrinsic values associated with aesthetic experience consist of meeting others outside the home and the possibility to break from routine, giving rise to relaxation. From a cultural-policy evaluation point of view, it is hardly relevant to measure the realization of these values because (1) they can be generated by any type of (outdoor) social activity, and (2) the

\(^2\) Note that, on the basis of the description of aesthetic experience in Chapter 6 of this research, the policy theory here is too positive, as the functions may not follow even although the values have been realized.

\(^3\) Note that for some types of personal development this is not entirely true. In researching the functioning of the performing arts in the social domain, some functions at personal level will be included in policy evaluation, namely, those that relate to skill development and personal efficacy or positive freedom (see Section 12.1.3).
functions they give rise to are more relevant to public policy. The functions at personal level in the social domain (being socially active, buying into the institutions, values and norms of the community) occur in the activity itself, which means that the sheer volume of attendance is indicative of the extent to which these functions occur. Note that this involves recording the number of visitors rather than visits to avoid overestimation. However, this is rarely recorded. Furthermore, data on the frequency of visits reveals the significance of the performing arts to the lifestyles of audience groups, which is in itself an indicator for the functioning of the performing arts in society.

In conclusion, it should be remarked that Personal Value creation can be evaluated through the realization of aesthetic values for audience members. Audience surveys can shed light on the realization of these values, but not on the occurrence of intrinsic functions. Furthermore, audience data (number of visitors, socio-economic background of visitors) should be used to assess the reach in society. Data on the frequency of visits to the (performing) arts can reveal information on the importance of the (performing) arts to the lifestyle of people.

12.1.3. Societal Value
In the economic domain, the effects of the performing arts come down to three issues:

- Direct and indirect employment effects.
- Attraction of knowledge workers as a result of a diversity of amenities.
- Influence on the image of a city or city district, which can lead to processes of gentrification.

Measuring the functioning of the performing arts in the economic domain thus relies on three types of research. The first type of research is impact analysis using type III multipliers and net multipliers. This is the most important method to measure arts’ functioning in the economic domain, both on the side of the extrinsic functions for artists (direct employment) and in relation to society (indirect employment). Second, the development of the number of knowledge workers in the city can be tracked when this is a stated goal of the cultural policy. Here, a problem of causation exists because (a) the relation is not one-sided (creative knowledge workers attract artists to the city as well as the other way around), and (b) knowledge workers will not be attracted by a tolerant climate and a diverse supply of performing-arts facilities alone. Other factors, such as horizontal career opportunities, the presence of venture capital and other leisure facilities such as sport, heritage and natural surroundings are important too. But when a city opts for it, it can commission research to monitor changes in the volume of knowledge workers. The third type of research is specifically relevant to the performing arts. As changes in real estate prices can be linked to the supply of performing arts in a city, multi-variate analysis of the difference in house prices between cities is an interesting way to evaluate performing arts policies. All three types of research cannot be restricted to single performances or institutions as they are done at an aggregate level: that of an entire industry in a city.
Chapter 9 related the functioning of performing arts in the social domain to the concept of social capital. Six measuring criteria for the functioning of performing arts in the social domain were identified:

- Skills development of attendees.
- Feeling well of attendees.
- Changes in personal efficacy (or positive freedom).
- Communal nature of the experience and sense of belonging.
- Organizational development (of predetermined communities).
- Changes in social capital within communities or the general population.

Note that the first four criteria actually concern the performing-arts’ functioning at personal level and therefore could also be included in the personal value part of the model for evaluation cultural policy. However, they specifically enable the transition from personal functioning to the functioning of the performing arts at societal level (in the social domain), which is why their inclusion here is more logical. Their evaluation should be based upon longitudinal audience research, which involves asking audience members how their aesthetic activities relate to their personal development. ‘Skills development’ denotes not only aesthetic skills but also skills in engaging with society (i.e., the skills to form bonds and bridges with other people) and organizational skills.

The development of organizational capacity within (predetermined) communities can be registered through observations. Changes in social capital can best be evaluated through the assessment of the social networks within communities (and society). Methods of doing so are discussed in Section 12.2.5. Note that there is a problem of causation here, as these effects of the performing arts cannot be causally linked solely to the aesthetic experiences of members of the communities under study. Therefore, it is more useful either to investigate the effects of the performing arts on pre-specified groups in society or to introduce an integrated evaluation of the outcomes of cultural, sports, welfare, education, and spatial planning policies.

12.1.4. Secondary Value Evaluation

As discussed in Chapter 11, secondary values concern the question as to whether or not the organization can effectively support the realization of primary values (its mission). Much of current policy evaluation effort focuses on these secondary values and therefore most measurement criteria and methods are readily available (see Tables 11.8, 11.9 and 11.10). From a public policy point of view, the relevant question is whether or not the arts organizations are able to effectively support the realization of aesthetic, personal and societal value on a continuous basis and for a variety of audiences.

Supporting realization of aesthetic value

Here, the artistic capabilities of the staff and the position of the art organization in the cultural field are relevant. This is reflected in the vocational education, artistic biographies of
the artists, and the budget for staff training. Engagement with amateur artists is also relevant to the realization of aesthetic value, for it contributes to the resilience of the cultural system in a city and it can create performances with specific characteristics that strictly professional performances may not have. Furthermore, the fiscal health of the organization is relevant as it reflects the long-term potency of the organization to offer its aesthetic services. However, this is only relevant when evaluating structurally funded organizations.

**Supporting realization of personal value**
Here, the ability of the arts organization to engage with its relevant audience is at stake. This involves marketing capacities (reflected in the education of support personnel and the budget for staff training) and marketing activities (which can be reflected in the funds devoted to marketing). Ancillary services are also relevant, most notably educational services, as they can enhance the realization of values for customers.

**Supporting realization of societal value**
This is the most difficult for single art organizations, as the realization of societal value typically relates to the total of cultural amenities in a city and is dependent on other than merely the aesthetic experiences of their audience members. Impact analysis (on an aggregate level) therefore seems most applicable. However, some measures at organizational level can also be conceived. The level of volunteer work expresses societal value, although not necessarily the value sought after in cultural policies. Furthermore, co-operation with organizations outside the cultural sector and the diversity of revenues of the organization reflect its ability to create societal value.\(^4\) Coverage of the organization in the general media is also a relevant indicator.

Note that, with the measures in these three areas, the tasks mentioned in cell A of Table 10.1 are investigated, as these involve the ability of the arts institutions to influence one another as well as amateur artists, and to engage with a variety of publics.\(^5\) Measuring the indicators for secondary value is relatively easy: they should be reported in the annual reports of the funded arts organizations (see Section 12.3.) Systematically gathering this data may considerably improve policy implementation because serious attention to these issues will drive the performance of performing-arts institutions in this respect, and thus stimulate their abilities to engage with their environment, which in itself can boost the functioning of the performing arts in society. Moreover, systematically following these issues should not only give insight into current levels of capacity - e.g., marketing skills measured through education levels of the staff - but also into changes in these levels and whether or not the capacities are actually used.

\(^4\) Note that the absolute amount of other subsidies and sponsor income are not suggested as indicators. Sponsors may have very different reasons for supporting arts organizations which are nothing to do with art’s functioning in society. Furthermore, it can be expected that sponsors will favour aesthetic activities that cater to the preferences of affluent groups in society, which means that sponsor income is not a good indicator. The diversity of sources of revenues is a better proxy for the ability of an organization to engage with its environment.
12.2. Measuring the Functioning of the Performing Arts in Society for Policy Evaluation

As can be discerned from the previous section, policy evaluation involves several measurement methods at various levels of aggregation. The measuring methods are discussed in this section.

12.2.1. Independent Advisory Boards

Independent advice is used frequently in cultural policy. In the model presented here, advisory boards can be used for assessing aesthetic value and for secondary value evaluation, although city administration itself has a role to play here too, as it gathers the necessary data from the annual reports of subsidized organizations. The question posed to the advisory board should be discussed here. As mentioned in Section 12.1.1, a retrospective approach is advocated for the evaluation of aesthetic value produced by the organizations. Apart from the intersubjective judgement of the quality of the productions, reviews of the work and prize money and/or awards are indicators that can be used.

Although artistic considerations are usually pre-eminent in their evaluations, advisory boards indeed evaluate management and marketing capabilities in some cases as well. Some may oppose the inclusion of these considerations. However, it is logical that the advisory boards – specifically for larger institutions – include secondary criteria in their evaluations in relation to the artistic profile of an arts organization and its position within the cultural system of the city. With the inclusion of ‘bestelcriteria’ (establishment criteria) by the Raad voor Cultuur and the functional quality criterion by the National Fund for the Stage Arts, current advisory board practices in the Netherlands are moving in this direction. Connected to this point is the issue of the composition of the advisory board. Its members should be able to evaluate independently both the aesthetic values and the management capabilities (when asked to do so). When both the city administrators and managers of institutions acknowledge the board members’ expertise in these areas, this improves the support for its advice. Therefore advisory boards can play a role in Aesthetic Value and in Secondary Value evaluation.

---

5 Therefore the model presented here includes what Holden (2004) denotes as ‘cultural systemic resilience’.
6 Under ‘functional quality’ the Fund uses a set of five criteria:
1. Contribution to diversity, which regards the contribution of the proposed productions to the diversity of the cultural supply; also ethnic diversity is meant by this.
2. Contribution to spread, which is regarded as geographic spread of the supply of the performing arts in the Netherlands.
3. Relationship between creation, production, programming and development of audiences, which is specific to the Dutch theatre system and involves the collaboration between producers and distributors of the performing arts.
4. Audience reach and development, which concern the question as to whether or not the applicant formulates a vision on the type of audience aimed at, and the development of such an audience in relation to the nature of the productions.
5. Cultural entrepreneurship and management, which concern the use of management techniques, guarantees for the continuity of the organization, and evidence of entrepreneurship by demonstrating that successful productions will be receive longer runs for instance (see NFPK+, 2010). Obviously, with the use of these criteria, the Fund veers into the realm of secondary value evaluation.
12.2.2. Audience and Reception Research

As the core of the policy legitimization concerns the values of aesthetic experience for the audience members, it is hardly conceivable to have an evaluation of cultural policy without surveying the performing-arts’ audience in the city. Audience surveys should provide data on the occurrence of intrinsic values at personal level and on the personal development of attendees, which is at the basis of the functioning of the performing arts in the social domain. Furthermore, audience research should shed light on the reach of the subsidized arts in urban society, and frequency of visits is an indicator of the significance of the performing arts in the lifestyles of people.

These last two issues can be researched quantitatively. This type of research is denoted as ‘audience research’ which aims at describing the features of audiences, e.g., demographic background, cultural habits and attitudes towards theatre and other leisure activities (Martin and Sauter, 1995, p. 27), or developing explanations for such patterns when the research takes on a theoretical-empirical approach (Schoenmakers and Tulloch, 2004, p. 18). The types of questionnaires needed are available. However, audience research does not shed light on the spectator’s experience in the theatre. This is the domain of reception research (Martin and Sauter, 1995, p. 29; Schoenmakers and Tulloch, 2004, p. 19). Reception research can include observations of the audience members’ reactions to performances (the easiest indicators here might be the duration of applause after a performance and laughter or intense silence during the performance). However, such observations do not shed light on the nature of the experiences and their significance in the lives of audience members. A number of methods have been developed ‘to retrieve information of a real theatre experience, though filtered through the consciousness, which verbalization necessitates’ (Martin and Sauter, 1995, p. 32). Martin and Sauter list methods such as semantic differentials, which involves asking audience members to choose between several adjectives without logical relation to the performance, non-directive probing interviews in which respondents are asked to comment on video-taped fragments of the performance they have attended, Theatre Talks in which a group of people freely chat about the experience, and questionnaires (ibid., pp. 32-3), although they do not describe the latter. For the purpose of policy evaluation, it is necessary to have focus groups in which a small group of audience members are interviewed about the experience, what it means to them, whether or not they feel they have been challenged by the performance, and whether or not they have felt this experience as a communal experience. If such focus groups are conducted longitudinally, they should be able to track the personal development of attendants and they can comment on the contribution of their (repeated) visits to the performing arts to this development, thus tracking the first stage of the functioning of performing arts in the social domain.7

Tulloch sees the use of these qualitative methods as a trend in cultural studies towards local, contextualized and situated knowledge production (Tulloch, 2005, p. 19). Although he
clearly does not dispense with these methods, as they generate important knowledge on the relationships between the aesthetic productions and their meaning to audiences in local circumstances, he thinks these methods are limited and that they represent a shift towards the research of alternative productions and venues, as they are smaller and therefore easier to study in depth. Therefore, he advocates the mixing of qualitative research methods with other methods to paint the full picture of the performing arts’ functioning in society. Qualitative analysis should therefore be supplemented with a quantitative research method. This is important to policy evaluation, as larger numbers of audience members can be included in the research, allowing more extensive conclusions on the functioning of the performing arts in society.

Schoenmakers and Tulloch (2004) also mention the lack of quantitative reception research methods, though they quote the example of Van der Blij (1995), already alluded to in Chapter 6 (see note 36). She asked audience members at ‘well-established’ performing-arts venues in the city of Groningen and also at venues that are only incidentally used for the performing arts, such as bars and community centres, whether they found descriptions such as ‘full of ideas’, ‘original’, ‘aesthetic’, ‘professional’, ‘convincing’, etc. applicable to the performance. These adjectives had been derived from the criteria used in the city’s arts policy. Boorsma and Van Maanen researched the experience of a play by the Noord Nederlands Toneel (North-Netherlands Theatre Company) performed in 2001. They developed a questionnaire with a series of general items based upon art theory (or rather: theories of art consumption) and a series of situation-specific items that reflected the artistic goals of the company and characteristics of the production (Boorsma and Van Maanen, 2003a, p. 166). For the present research, the items relating to art theory are most relevant. For instance, respondents were asked to either agree or disagree with statements that the performance ‘is thrilling theatre’, ‘is innovative theatre’, ‘is experimental theatre’, ‘is art’, ‘challenges one to see things differently’, ‘is touching’, ‘is confronting’. Furthermore, they included items related to entertainment and social interaction: the performance ‘is entertaining’, ‘provides a nice/sociable evening’, ‘offers material to talk about with others’ (ibid., pp. 168-70). Although their aim was to identify which items are most relevant to the overall evaluation of the experience and not to the artistic nature of it, their research is certainly interesting, as several items reflect values identified in this research. These examples demonstrate that quantitative

Note that the perceived communality of the experience is influenced by the organizational setting of the performing arts. Specifically when the communal nature of the experience is organized, such as is frequently the case for amateur and community arts, focus group interviews can yield much information. Her results corroborate the fact that aesthetic experiences in commercial and amateur arts can yield the same values as those in the subsidized arts, a position taken in the present research as a consequence of pragmatist aesthetics (see Chapter 6). However, her adjectives have been derived from policy documents and reports by independent expert advisors, which raises the question as to whether or not they indeed relate to the artistic nature of aesthetic experience. It seems rather that they relate to artistic quality as a property of productions, now being applied by lay audience members. This severely damages her claim to the equality of the values experienced in both settings.

See also Boorsma and Van Maanen (2003b) where they reported on the same research. Here their aim was to identify the influence or reviews on the experience.
methods can be used to shed light on the nature of aesthetic experiences, i.e., they provide qualitative information.

For the present research a recent study by Wilders is of specific interest as she states that some of the aesthetic values identified in the present research can be found in the questionnaire she used to investigate the influence of theatre buildings on reception (Wilders, 2009, p. 496). She developed a questionnaire ‘to measure the outcome of a performance based on the experiences of individual recipients. The [questionnaire] (...) offers an example of how the theatrical experience and the possible values realized through this experience can be made operational and thus measurable’ (ibid., p 461). In 2006-2007, she conducted an audience survey among the audiences of several productions performed in a number of theatres across the Netherlands to determine which values the audience members derived from the experience. In Table 12.1, (part of) the items in her questionnaire are related to the values of aesthetic experience used in the present research.

As can be gathered from Table 12.1, her questionnaire items correspond to the values as identified in this research. However, a few problems with the manner in which these values were made operational remain. As Wilders did not distinguish between non-artistically aesthetic and artistically aesthetic values, the wording of the questionnaire items concerning the use of fantasy or imagination and relaxation or entertainment does not lead to clear conclusions on the artistic nature of the experience. Essentially the questionnaire lacks questions concerning whether or not the experience has been challenging. None the less, it is a promising conclusion that audience members were able to reflect on the nature of their experience in the items suggested by Wilders. Although the questionnaire was not designed to record the values (and functions) of aesthetic experience as they are presented in this research, and only in hindsight appear to do so, it should be possible to modify it to indeed measure the occurrence of intrinsic values of aesthetic experience.

Therefore, it can be concluded that, by means of audience and reception research comprising both focus groups and questionnaires, it is possible to assess the reach of aesthetic services in the urban society, to track the occurrence of intrinsic values of the aesthetic experience, and

---

10 Wilders here refers to a first presentation of these values in Van den Hoogen (2009).
11 Her research aim is to investigate whether theatre architecture influences the experience of performances. The experience of the architecture of the venues is not relevant to the present research. The discussion of her research is based upon the research items that measure the theatrical experience, not the experience of the building. Note that in researching the performances of five productions (three plays and two modern dance performances) in different venues she was able to control for the aesthetic values that are present in the performances. Thus she avoids a pitfall Schoenmakers and Tulloch identify for audience research: a need to include an analysis of the theatrical work itself (Schoenmakers and Tulloch, 2004, p. 19).
12 Wilders claims not to have included the functions of aesthetic experience in her questionnaire as they ‘are more likely to become apparent over time and to be the result of a combination of [aesthetic] experiences, and (...) therefore require a different methodology’ (Wilders, 2009, p. 469). Upon closer examination, one questionnaire item, ‘The performance reconfirmed my understanding of the world around me’, clearly refers to the artistic-aesthetic function of ‘possible change of views and insights’ although the wording of the item only represents one of the possible reactions to the challenging of perceptions, namely, where the audience member chooses not to alter his or her views.
### Values and functions of aesthetic experience as identified in this research

<table>
<thead>
<tr>
<th>Intrinsic values</th>
<th>Items in the audience survey (Wilders, 2009)*</th>
</tr>
</thead>
</table>
| **Excitement due to the experience of non-present worlds** | I thought the scenery (lighting, set design, costumes, etc.) was beautiful  
I found the behaviour of the characters fascinating / I found the aura of the dancers fascinating  
I was affected by the style of play/dancing  
I was impressed by the way the actors used the space  
I was totally absorbed in the world of the performance  
I felt the physical tension transmitted by the actors/dancers  
I had the feeling I was truly addressed by the actors/dancers  
I had the feeling I truly connected with the actors/dancers  
I found the performance relaxing (‘een fijne ontspanning’)**  
I found the performance entertaining (‘vermakelijk’) |
| **Empathizing with imagined emotions** | The development in the plot/choreography appealed to me  
I thought the performance had a satisfactory ending  
The theme appealed to me  
The way the theme was dealt with appealed to me  
I thought the scenery (lighting, set design, costumes, etc.) related well to the theme  
I thought the acting/dancing was good  
I found the performance moving  
(I found the performance tedious)  
I empathized with the emotions expressed  
I could identify with one or more characters/dancers |
| **Delight in the use of the power of imagination** | I found the performance surprising  
I found the performance confrontational  
I found the performance difficult  
(I found the performance tedious)  
(I found the performance not very innovative) |
| **Experience of new perceptions** | I found the performance moving  
I found the performance inspiring  
The performance triggered my imagination  
The performance made a profound impression  
The performance struck a chord  
(The performance did not add anything to my own experiences) |
| **Testing one’s views and insights** | The performance offered me a different view on things  
The performance offered me an alternative view of my own life  
The performance offered me insight into the human condition  
The performance challenged my outlook on things |
12. Policy Evaluation using the Framework for the Functioning of the Performing Arts in Society

<table>
<thead>
<tr>
<th>Intrinsic functions</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sublimation of needs and satisfaction of sublimated needs</td>
<td></td>
</tr>
<tr>
<td>Change or reconfirmation of views and insights</td>
<td>The performance reconfirmed my understanding of the world</td>
</tr>
<tr>
<td></td>
<td>around me</td>
</tr>
</tbody>
</table>

* Items relating to the experience of the theatre venue have not been included in this table.

** Including this item here is difficult. By stating ‘een fijne ontspanning’ (relaxing), Wilders clearly does not aim at measuring ‘intellectual entertainment’.

Table 12.1 Wilders’ questionnaire related to intrinsic values and functions of aesthetic experience.

to assess the first stage of the functioning of the performing arts in the social domain. Note that these techniques can provide institutions with information as to which intrinsic values they generate. Therefore, they might be able to correct their artistic policies according to such data. However, the ability of programming institutions to do so is limited, as they do not decide on the aesthetic properties of the productions staged. None the less, these surveys are as useful to them as to the producing companies, because the programmers have the responsibility to organize the realization of aesthetic values for their audiences. Their role in organizing the right public for the performances and the best possible conditions under which the realization of the values can take place can also be supported by information on the realization of these values. Furthermore, the reception research can generate information on the value of supportive services and provide useful marketing information. Because conducting these surveys can be a costly matter, city administrators should provide the funds for conducting them by allowing for sufficient overhead costs

12.2.3. General-Public Surveys

As it cannot be argued that aesthetic experiences in the non-subsidized sector do not yield the intrinsic values and functions that are the focus of cultural policy, policy evaluation should include a more general picture than the subsidized sector alone. The use of all cultural facilities in a city can be assessed by means of general-public surveys. Researching this item involves including, in regular populations surveys, questions on the use of cultural facilities in the city in the last twelve months. Questions can also cover the frequency of use and when questions also concern the socio-demographic background of respondents, a picture of the aesthetic activities of the whole population can be painted, including the particular aesthetic services that appeal to particular social strata. It is not necessary to conduct such research on an annual basis, as aesthetic consumption patterns are generally reasonably constant (see Martin and Sauter, 1995, p. 28). This picture is needed to assess the reach of aesthetic services in society. As demonstrated, some of the cities in the research sample use the general-public survey to gather such data. It seems possible, at least in theory, to conduct the same reception research as presented in the previous section in order to gain insight into the nature of the aesthetic experiences in the non-subsidized settings. In
Part IV: Policy Evaluation

practice, however, this seems hardly possible. It can be useful when questions arise as to the position of specific groups in society.

Including specific questions in general public surveys might also yield evidence of the functioning of performing arts in the social domain. Van den Broek has used surveys of ethnic minorities (Turks, Moroccans and Surinamese) and compared them to data on the general Dutch population. His aim was to measure the contribution of cultural activities to social cohesion, which he made operational through five indicators: identification with (Dutch) society, trust, respect for others, and contacts with other ethnic groups during leisure activities. He regards these indicators as measures for bridging social capital. The parameters were measured by asking respondents to general-public surveys about the extent to which they feel Turkish, Moroccan, or Surinamese rather than Dutch (5-point scale), whether or not they trust people in general (yes or no), how they feel towards other ethnic groups (expressed on a ‘thermometer’ scale from 0 to 100 degrees), the extent to which they agree with the statement that most Dutch people respect Muslim culture and most Muslims respect Dutch culture (5-point scale) and the extent to which they meet native Dutch and other ethnic groups during leisure time (often, sometimes, never) (see Van den Broek, 2009, p. II.1-3-25). A link between participation in cultural events and social cohesion can be established specifically for the inter-ethnic leisure contact. However, this cannot be considered as a causal link, as a certain openness of mind is needed for both types of activities. Van den Broek suggests that openness of mind is the underlying variable explaining the concurrence of cultural participation and inter-ethnic contacts. For ethnic groups, cultural participation cannot be linked to trust in and respect for other ethnic groups (ibid., p. III.1-3-27/28). This leads Van den Broek to conclude that there is no empirical evidence for a policy theory that assumes that cultural participation will increase social cohesion. However, given the fact that he himself suggests that openness of mind is a common explanatory factor of cultural participation and social cohesion, he might have been less strict in this conclusion. Research into the contribution of cultural participation in developing this openness of mind is certainly necessary. However, this type of research cannot be based upon information from general-public surveys but must be gained from in-depth interviews of participants in cultural events. None the less, Van den Broek is right in concluding that the link between cultural participation and social cohesion is not one-sided and cultural participation in itself can be an expression of social cohesion.

12.2.4. Economic Impact Analysis

Methods to research the direct and indirect economic impact of cultural facilities in a city have been discussed in Chapter 8. Impact research (using type III net multipliers) should be commissioned to investigate this. This research starts with the salaries paid and expenditures on the procurement of materials by cultural institutions, which are stated in their annual reports. In the research period, none of the cities in the sample commissioned economic impact analyses. Note that the earlier researches on economic impact in Groningen (Julien, 1989, and Julien et al., 1997) and Amsterdam (Hietbrink et al., 1985, and KPMG, 2006a, b and
c) do not use the correct multipliers. This means that, to date, there are no correct assessments of the economic impact of culture in the Netherlands apart from the multivariate analysis of real estate values (Marlet, 2009). Furthermore, cities can opt to research changes in the proportion of knowledge workers in the city.

12.2.5. Measurement of Social Capital
Apart from the questions discussed in the section on reception research, there are specific methods to measure changes in social capital, although measurement is no easy task (Field, 2008, p. 143). Measuring social capital within a macro-perspective involves assessing the total amount of social capital present within a community, and assessing changes through longitudinal research. Several authors discuss the problematic of measurement of social capital in the macro-perspective (see Flap, 2004; Van der Gaag and Snijders, 2004; Völker, 2005; and Field, 2008). They conclude that there is a lack of longitudinal data and uncertainty as to the specific indicators that should be included in the research. Flap (2004) therefore suggests that a micro-perspective may be a more promising route for measurement. As the rewards of social capital are causally linked to social relationships, the assessment of individual social capital seems the most valid way to evaluate the impact of the performing arts in the social domain. This perspective also allows for a focus on the question as to the conditions under which social capital is generated. This is important as the origin of social relations and networks is related to characteristics of the individual and social contexts (Völker, 2005, p. 14). Theoretically one could ‘map’ the social relations of all members of society and thus achieve a total map of the social capital. However, Völker’s remark implies that research of the effects of the experience of the performing arts on social capital will be most successful when the research focuses on pre-specified groups whose social conditions can be known beforehand. When it is possible to assess changes in the social capital of specific individuals or communities, at least theoretically, the effects of the experience of the performing arts on social capital can be monitored, although there will be a problem of causality, as changes in the social networks of individuals will be the result of many more influences than only aesthetic experience. Specifically for cultural policy, there is a link between bridging ties and the artistic nature of aesthetic experience. The measurement therefore should focus specifically on these ties.

Measuring social capital at individual level involves charting the social network of an individual and the resources that this network gives access to. However, this is not without problems because the resources that can be accessed are goal-specific. It depends on an individual’s goals when considering the resources he or she will access through his or her network. For the purpose of cultural policy evaluation, the goals cannot be specified beforehand. This means that the measurement instrument should cover a wide variety of different social resources (Van der Gaag and Snijders, 2004, p. 201). Furthermore, individuals

---

13 Putnam, for instance, used no less than 14 indicators to make up a social capital index (see Field, 2008, p. 144).
14 Although with the introduction of more sophisticated statistical techniques (see Sabatini, 2009), such problems may be overcome in the future, which would mean that data from general-public surveys will become important in assessing changes in social capital.
differ in their command of personal and institutional resources, which leads them to make
different choices when deciding about using social capital.\textsuperscript{15} These differences necessitate
measuring potentially available resources in social networks rather than the resources one
actually has accessed through social networks (\textit{ibid.}, p. 203). Methods to measure individual
social capital are discussed by Van der Gaag (2005). These networks are assumed to exist
between the respondent (ego) and others (alters), therefore they are called ‘ego-centred social
networks’. They can only be studied by observation in limited settings (e.g., classrooms or companies). In most studies on social capital in the general population, individual respondents mention their network in an interview situation. Highly structured and
systematic queries are used. Van der Gaag discusses three interview techniques: the Name Generator, the Position Generator, and the Resource Generator.\textsuperscript{16}

The \textit{Name Generator} asks respondents to identify social network members. Several types of
Name Generators have been designed, varying in the way the respondent is reminded of the
possible existence of social ties. For instance, the interviewer may ask about family relations
or affective relations. Van der Gaag, however, favours the exchange relationship generator in
which respondents are asked to identify people with whom they have performed, or may in
the future perform, various specific types of support exchanges. The relationships thus
traced are goal-specific and provide information on the social resources accessible through
the relations (\textit{ibid.}, p. 80). In a second phase of the interview, which is called the ‘interpreting’
phase, the respondent is asked to score all of the relations generated on several
characteristics, such as duration of the contact, frequency of the contact, physical proximity
(lives under or over 5 km away), perceived intensity of the contact and perceived trust (both
measured on a five point scale). Tie strength can be assessed by means of such interpretation.
To a large extent, the Name Generator relies on the ability of the respondent to remember all
his or her social contacts. This is a problem when it comes to weak ties, therefore the Name
Generator seems less suitable for the present research. Also, differences in interpretation
between interviewer and respondent limit comparability of the results. Furthermore, much
information is redundant because many alters will provide access to the same resources. The
instrument can best be used to indicate the size of social networks, rather than the diversity
of resources. A final downside of the Name generating and interpreting technique is its time-
consuming nature. It is a heavy burden for both interviewer and interviewee to complete the
interview especially when larger networks are encountered (\textit{ibid.}, p. 137).

\textsuperscript{15} Individual capital refers to the abilities of the person him- or herself, mostly acquired through education, and
institutional capital refers to resources becoming available through social institutions (government agencies, one’s
place of work) with whom one does not maintain personal relationships (Van der Gaag and Snijders, 2004,

\textsuperscript{16} Note that the techniques discussed do not identify existing social ties between different alters. Investigating
the structure of social networks would imply also charting these relationships. The methods discussed here however
can be used to composite proxies for network density (defined as the amount of ties between all actors in a social
network). However, for the present research network density in itself is not relevant. The access to social
resources gained through the network is what matters.
The *Position Generator* is an instrument in which respondents are asked to indicate whether or not they know an acquaintance, friend or family member with a certain occupation. The assumption is that occupations with higher prestige will make more resources available to the respondent. Furthermore, it is assumed that the strength of the tie between the respondent (ego) and the person with the occupation (alter) determines the likelihood that alter will make the resources available to ego. Thus, an individual’s amount of social capital can be assessed. It should be remarked that the instrument involves a ranking of occupations chosen from regular census studies of the population in question. The esteem attributed to occupations will differ from (sub)culture to (sub)culture. The Position Generator can be used to assess the distribution of social capital in society. However, this involves comparison of data from different studies in which it is crucial that the same occupations have been used and are worded in the same way. The Position Generator has a few weaknesses. For instance, respondents may not know the occupation of all their network members. Therefore the Position Generator will yield an underestimation of the occupations accessed by the respondent. Furthermore, the wording of the occupations is crucial. In general, the more specifically an occupation is worded, the smaller the chance of a positive answer. Therefore, for comparison purposes, the wording needs to be exactly the same. Furthermore, the Position Generator can yield false reporting of social capital, as some occupations, such as doctors, clergymen and teachers, may be known readily, although it is questionable whether or not they would make their social resources available to everyone they know.\(^\text{17}\) Finally, for questions regarding health and general well-being, the Position Generator is not the most useful instrument because it leaves emotional support, companionship, and help in times of illness out of the equation. The Position Generator may not be able to reflect the actual network size, because the total number of relations is limited to the total number of occupations suggested to the respondent, although several other measures of the social network of respondents can be constructed. According to Van der Gaag the most obvious measure to use is the number of positions accessed, for this can be considered a proxy for the diversity of resources accessed by the respondent (*ibid.*, p. 131). However, the assumption behind the Position Generator is that specifically higher-prestige occupations give access to more social resources and that lower-prestige occupations give access to different kinds of resources than the higher-prestige occupations. This implies that the logical measure is a combination of the highest accessed prestige and the range of prestige accessed.\(^\text{18}\)

The *Resource Generator* combines the characteristics of both the Name Generator and Position Generator in asking about access to a fixed list of resources ‘each representing a vivid,

\(^{17}\) ‘Knowing someone’ was made operational by asking respondents if, on accidentally meeting the alter on the street, they would recognize him or her, would know his or her name, and could start a conversation with him or her (Van der Gaag, 2005, p. 104). This, however, does not entirely solve the problem of false reporting.

\(^{18}\) Van der Gaag makes yet another important remark about the Position Generator. Although the measures taken from the Position Generator have a clear theoretical basis, they contain only indirect information about the social capital accessed. Furthermore, the concept behind the position generator hinges on job prestige and therefore the measurement instrument is slanted towards researching goal-specific social capital, i.e. weak ties (Van der Gaag, 2005, pp. 137-8). For measuring the impact of performing arts in the social domain, this is not a drawback of the Position Generator.
Part IV: Policy Evaluation

concrete subcollection of social capital, together covering several domains of life’ (Van der Gaag, 2005, p. 138). The questionnaire uses the same indicators for the strength of the ties as the Position Generator; i.e., the respondent is asked to indicate whether the resource is available through an acquaintance, a friend or a family member. The selection of items in the questionnaire should result from systematic theoretical considerations about which social resources represent the ‘general’ social capital of individuals in a population (ibid., p. 138). In an inquiry into the social capital of the Dutch population (performed in 1999/2000), a list of 33 resources was used, ranging from ‘someone to babysit’, ‘someone who owns a car’, ‘someone who knows people in politics’ to ‘someone who can help in finding a summer job’. The Resource Generator is able to distinguish between types of social capital that relate to differences in the goals of individuals. This feature makes it suitable for researching resources that cover a variety of sub-domains of life (Van der Gaag and Snijders, 2004, p. 208). The Resource Generator enables better assessment of the size of social networks, as respondents are asked to identify more than one person they know from whom they could acquire the suggested resource. Nevertheless, there is still a restriction here because the number is limited to a maximum of 5 per resource. Furthermore, the measurement instrument allows the combination of an assessment of network size and types of relations with information on the resources accessed through the network. None the less, several problems with the Resource Generator remain. First, what if the respondent is able to provide for the resource himself or herself, or pay someone to help him or her? This implies that when longitudinal studies are used to track changes in social capital (either of general populations or of specific groups within the population), the results should be controlled for education and income levels. Second, the instrument does not measure the willingness of alters to make the recourse available to the respondent. The assumption is that the strength of the ties is a proxy for the willingness of alters to make the resource available. Third, the instrument is difficult to construct as it involves choosing the items to be included in the questionnaire. The relevancy of certain items may differ widely for respondents. Therefore, the instrument might be especially useful for researching specific populations rather than the general population. For the present research, this does not necessarily present a drawback.

Van der Gaag advocates measuring individual cultural capital through a combination of instruments, as each instrument seems to measure different types of social capital. A combination of the Resource and Position Generator is easy to implement because both methods can be included in one questionnaire and thus are least costly to investigate (ibid., p. 202). Furthermore, both types of research are adaptable to the type of social capital the researcher is interested in, as the positions or resources suggested to the respondent can be manipulated. For the present research, the most important parameters of social capital – in order of significance – are:

---

19 Note that improvements in income and education levels may also be a result of changes in personal lives that are trigged by the social capital the individual has accessed. Therefore these changes in themselves may also be a result of the policies directed towards the individuals concerned.
• The amount of weak ties within a person’s network, because it is especially through weak ties that resources that are important to improve one’s socio-economic position can be accessed – which is the ultimate goal in social policies – and also because weak ties in particular can be connected to artistically aesthetic experience.

• Diversity of the contacts in the network as a proxy for the diversity of social resources accessed through these contacts.

• Network size as a proxy for the ability of the individual to access specifically bonding relationships.

Changes over time in these variables should be measured in order to evaluate the effect of policies. This implies longitudinal research using exactly the same questionnaires. Furthermore, the changes should be corrected for the development of the personal capital of the respondents which can be monitored through income and education levels.

In Table 12.2, the measurement methods discussed by Van der Gaag are linked to the above-mentioned measurement parameters. The bias of the Name Generator towards strong ties is a specific drawback for the present research. Therefore, the use of a combination of the Position and Resource Generators seems most applicable. However, these techniques still have several weaknesses which need to be addressed in further research aimed at developing measurement instruments for social capital. The following remarks must be made.

First, the most important issue concerns the question as to whether or not participation in the (performing) arts leads to specific forms of social capital, and whether such forms can be measured. The theories presented in Chapter 9 and the measurement methods presented in this section cannot shed light on this. Further research is necessary here, although it seems that a combination of the quantitative measurement instrument presented here and the qualitative longitudinal focus groups mentioned in the section on audience and reception research is most applicable.

Second, it seems that, as the assessment of the social capital of a general population is difficult, more insight can be created by researching specific groups in society as the instruments can be geared specifically to the social resources needed and accessed by these groups. For policy evaluation purposes, focusing on specific groups is not problematic. It seems useful to carry out research tailored to these groups, certainly when specific groups are targeted in the policy. Van der Gaag warns, however, that general differences in the distribution of social capital in a population can distort or obscure the results of studies on the social capital of specific groups. In the Dutch general population, it appears that social capital is distributed unevenly: better-educated people and men tend to have access to more diverse and higher-prestige occupations (ibid., p. 130).
<table>
<thead>
<tr>
<th>Weak ties</th>
<th>Name Generator</th>
<th>Position Generator</th>
<th>Resource Generator</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>--</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>(Typically weak ties will not be remembered by the respondent easily, so the NG is biased towards strong ties)</td>
<td>(The nature of the contact is researched through asking whether the contact is an acquaintance, friend or family member)</td>
<td>(The nature of the contact is researched through asking whether the contact is an acquaintance, friend or family member)</td>
<td></td>
</tr>
<tr>
<td>Diversity of contacts</td>
<td>+</td>
<td>±</td>
<td>+</td>
</tr>
<tr>
<td>(Can be assessed in the interpreting phase)</td>
<td>(Assessed through the range of prestige of contacts of the respondents)</td>
<td>(The resources suggested to the respondent should be diverse, covering several domains of life)</td>
<td></td>
</tr>
<tr>
<td>Network size</td>
<td>+</td>
<td>-</td>
<td>±</td>
</tr>
<tr>
<td>(Is assessed through the range of positions accessed. The total number of contacts is restricted to the total number of positions suggested to the respondent and thus total network size is not measured in itself.)</td>
<td></td>
<td>(Assessment of network size is limited as a maximum to the amount of contacts per resource is introduced)</td>
<td></td>
</tr>
<tr>
<td>Drawbacks of the instrument for the present research</td>
<td>The NG is costly in execution, especially when encountering large networks. The NG is biased towards strong ties.</td>
<td>The PG is least able to assess network size.</td>
<td>The RG is not easy to construct as it involves choosing which resources to include in the questionnaire.</td>
</tr>
</tbody>
</table>

Table 12.2 Characteristics of measurement instruments of individual social capital for evaluating arts policy

Third, the question when to measure the effects of cultural policy remains a difficult one, at personal as well as societal level. In the social domain, the question needs to be answered concerning the moment when changes in social capital (of specific groups) may be expected to have occurred as a result of the policies. From a scientific point of view, this question cannot be answered. The time frame for research therefore becomes a political decision rather than a scientific one.
Fourth, these measurement instruments do not remedy the drawback that the causal relationship with aesthetic experience is weak because many other factors influence changes in cultural capital. Using these methods therefore is only logical in two instances:

- When one wants to track the effects of specific cultural policy measures for specific individuals or groups in society. This is only possible for participants in specific cultural interventions, such as the community arts.
- When one wants to track the effects of the combined efforts of cultural, sports, welfare, education and spatial planning policies on social capital. Researching the combined effect of these policies has the advantage of enabling more costly types of research and it overcomes the problems of causality. The drawback is that it will not be able to evaluate individual policy measures. Therefore it is not possible to identify which interventions are most effective. A further advantage of this type of combined effort might be that the concept of social capital may lead to alignment in these different policy areas, each with its own dynamics and, in many cases, implemented by different branches of the municipal organization.

12.3. Organizing Art Policy Evaluation

12.3.1. Cultural Institutions and City Government
Arts policy evaluation involves a joint effort by the city administration and the cultural institutions. A large part of the information needed is reported (or should be) in the annual reports of cultural institutions. Furthermore audience and reception researches involve the arts institutions as much as the city administration. In Figure 12.2, the necessary research activities and data to be included in the annual reports is listed in each of the fields for policy evaluation.

_Aesthetic value evaluation_ can be delegated to independent advisory boards. The city administration should formulate the questions to be put to the board. This involves deciding on the extent to which management and marketing criteria (and thus secondary evaluation criteria) should be included in the independent advice. Furthermore, the city administration should decide on the composition of the board. When city administrators and the management of cultural institutions both acknowledge the expertise of the members of the boards regarding artistic, and management and marketing issues, the acceptance of the advice will be strengthened. Therefore, the city administration should consult the arts organizations before appointing the members of the board. The annual reports of the organizations should include the data needed for evaluating aesthetic value (and management and marketing capabilities). As the city administration gathers the annual reports, it is logical that it will supply this data to the board.

_Personal value evaluation_ requires joint effort by arts organizations and the city administration. The arts organizations, as they have access to the audiences, should conduct audience research to provide the data on reach of cultural facilities and the socio-demographic composition of the audience. Furthermore, the quantitative questionnaires that provide
information on the occurrence of intrinsic values should be conducted by the arts organizations. These audience surveys will provide the organization with valuable information to inform management and marketing strategies.\textsuperscript{20} Note that, as cultural consumption patterns are fairly constant, such research need not be conducted on an annual basis. The city administration should provide the funds for the audience surveys by allowing for sufficient overhead costs. It is up to the city administration to aggregate the quantitative information on the reach of the aesthetic services, and to report on this aggregated information to the city council and the general public. This data should be supplemented with the information on the reach of non-subsidized cultural presentations in the city, which can be obtained through general-public surveys. Furthermore, the city administration should conduct (or commission) the longitudinal qualitative reception research (focus groups) needed to complement the quantitative data on the occurrence of intrinsic values.

\textit{Societal value evaluation} is totally up to the city administration, except that the annual reports of the arts organizations funded should provide data on the salaries paid in order to assess direct employment. The city administration should commission the relevant research to analyse economic and social impact as both types of research involve expert knowledge of the economic and social sciences. For the social impact analysis, the longitudinal reception research involving focus groups should shed light on the importance of aesthetic experience to the personal development of audience members, which is relevant for changes in social capital.

Finally, \textit{secondary value evaluation} can be based upon information which is readily available in the annual reports of subsidized institutions. The city administration and the arts organizations should enter into an agreement as to what information should be included in the annual reports. Such agreements should be made for longer periods of time and not be altered during the process, as this will not encourage adherence to the agreement by the funded institutions.\textsuperscript{21} As stated above, secondary value evaluation can in part be delegated to independent advisory boards when they include management and marketing criteria in their judgements.

Thus the cultural institutions and the city administration have a shared responsibility for cultural policy evaluation. However, as it is the exclusive responsibility of the administration to report on the execution of policies to the city council (and a wider public), whereas coordinating the evaluation effort is up to the administrators.

\begin{flushleft}\textsuperscript{20} Note that the assumption of the autonomy of artists presupposes that it is the arts organizations that should decide on the consequences of the outcomes of research on personal value realization for their artistic, management and marketing strategies. This presents a further reason why the arts organizations should conduct such research themselves (see Section 12.4.1).
\end{flushleft}

\begin{flushleft}\textsuperscript{21} Note that the evaluation model presented here presupposes that managers of cultural institutions are able to formulate their missions on the basis of notions on the societal functioning of the arts. This may be regarded as a reflection of the growing professionalization of the arts sector. However, as discussed in Chapter 11, managers may lack the capacity to formulate a mission statement in this respect. City officials need to reckon with this and support managers of publicly funded institutions.
\end{flushleft}
### Primary Value Evaluation

#### Aesthetic Value

**Assessment of artistic quality of productions**

**Annual reports:**
- Critical and peer reviews
- Prize money and awards
- Co-operation with other artists
- (Professional) media coverage
- Percentage of funds devoted to risk full productions
- Commissions of new works

**City Administration:**
- Formulate questions for independent advisory board
- Decide on composition of advisory boards

#### Personal Value

**Assessing the occurrence of values and functions at personal level (and at the level of the audience as a community)**

**Cultural institutions:**
- Quantitative reception research

**City Administration:**
- Provide funding for audience research
- Longitudinal qualitative reception research (focus groups)

**Assessing the reach of aesthetic services in the city**

**Annual reports:**
- Number of visitors
- Frequency of visits
- Socio-demographic background of audiences

**City Administration:**
- Aggregation of audience data from annual reports
- General public surveys on use of cultural amenities

### Secondary Value Evaluation

**Assessing the ability of performing-arts organisations to deliver aesthetic services on a continuous basis to a variety of audiences:** annual reports should include information on:

- Vocational training of artistic staff and artistic biography
- Staff training budget
- Activities with amateur artists
- Fiscal health

- Marketing training of the support staff
- Staff training budget
- Funds devoted to marketing

- Educational services
- Funds devoted to educational services

- Number of volunteers
- Co-operation with organizations outside the cultural sector
- Diversity of revenues
- General media coverage of the artists/organization

### Societal Value

**Economic Impact Analysis**

**Annual reports:**
- Salaries paid to staff
- Procurement of materials

**City Administration:**
- Research through existing impact techniques
- Research changes in the proportion of knowledge workers in the city (optional)
- Multi-variate analysis of changes in real estate values

**Social Impact Analysis**

**City Administration:**
- Longitudinal reception research (focus groups)
- Research on changes in social capital (can only be done for specific groups in society)

---

**Fig 12.2** Arts-Policy Evaluation Using the Framework for Describing the Functioning of the Performing Arts in Society (actions to be done by city administration and subsidized institutions, and information to be included in annual reports of cultural institutions)
Of the eight cities in this research, Groningen and Rotterdam are evidently the cities with the most elaborate evaluation instruments. Thus, it appears that – just as is the case with arts policy formulation itself – municipal arts-policy evaluation develops in stages. This implies that cities can choose how to evaluate arts policy. Here political considerations will collide with administrative concerns. Politicians might opt for extensive evaluation mechanisms to meet the challenges of the evidence-based policy era, although they also might opt differently and advocate low bureaucracy and simple and swift government. Bureaucrats may view evaluation mechanisms as a way to account for spending of public money and ‘defending’ the cultural budgets within the city administration. But their interests in policy evaluation can be more sophisticated. For one, these techniques presuppose agreement between the administration and cultural institutions on the data that will be collected in the course of policy execution, thus facilitating the relationship between institutions and the city administration. The model for policy evaluation can also be used to align cultural, sports, welfare, education and spatial planning policies. Thus, both city administrators and politicians should find reasons enough to move from one stage of policy evaluation to the next.

12.3.2. Stages in Arts Policy Evaluation

**Stage 1: Advocacy**

Key Question: Do the cultural institutions in the city provide a cultural presentation with artistic quality?

In its most rudimentary form, arts-policy evaluation relies on the framework presented in Figure 10.1 and it assumes that, when artistic quality is produced in the city and experienced by citizens, the values and functions associated with aesthetic experience will ensue. This type of evaluation practice can only be characterized as advocacy. It relies on the story the framework tells to legitimize policy efforts. Policy evaluation will be restricted to counting the number of performances and the attendances at these performances. The role of independent advisory boards with experts on artistic quality (as a property of the productions realized) is crucial. Their evaluations will drive policy. Furthermore, the city administration can research the reach of the productions through audience data gathered by institutions.

Though it might seem at first glance that all cities comply with this stage by making use of independent advisory boards, this stage in itself can be a step up from current policy evaluation practices. First, this stage involves systematic use of advisory boards for all institutions in the city (both producing and distributing organizations). Second, it involves retrospective evaluations. This is not the case in all cities. Note that cities using this stage of cultural policy evaluation will be vulnerable to the proponents of the accountability movement, although, when they can demonstrate a structured approach to the evaluation of
artistic quality and research on reach, this may be considered a step up from a mere handing out of public funds to artists.

**Stage 2: Cultural Governance**

Key Question: Are the cultural institutions in the city capable of delivering the aesthetic presentation that provides the benefits associated with art to a variety of citizens in the long run?

The second stage implies that the city administration will supplement the policy evaluation with the secondary value evaluation techniques. This means that the administration chooses to ensure that the benefits of aesthetic activities in the city will be there over a longer period of time. This means that they need to ensure that the cultural institutions are well-run institutions which can deliver these benefits to a variety of citizens over a long period of time. In this stage, the annual reports of the cultural institutions should be more elaborate and include measures that reflect the marketing and management capabilities of the organizations, and more elaborate financial controls can be installed to insure fiscal health over a longer period of time. Adhering to standards within the cultural industry becomes important, through benchmarking and adherence to cultural governance codes for instance. It appears that this stage is often forgotten in policy evaluation because it is not a very visible stage. Adherence to such codes of cultural governance can be checked by the administration and is not usually the subject of heated political debates. The introduction of cultural governance codes in the Netherlands is, however, fairly recent; therefore this stage is a step up from current practices.

**Stage 3: Occurrence of intrinsic values and functions for audience members**

Key Question: Do the cultural institutions in the city provide the personal benefits associated with art to a variety of citizens?

In the third stage, the policy evaluation techniques are complemented with interest in the reception of the performing arts. Where, in the second stage, marketing data on the audience will be collected as a basis for formulating institution marketing policies, audience surveys in this third stage should also include methods to evaluate the nature of the experiences afforded to the audiences. Again, it should be noted that policy evaluation in this stage is still dependent on an advocacy point of view. The assumption here is that, when intrinsic values and functions at personal level are realized and it is ensured that cultural institutions are able to do so on a long term basis, the societal values and functions will follow. But here the advocacy is backed up by real data on the occurrence of personal values. Thus, the core of the policy legitimization is actually included in the policy evaluation. In this stage, the evaluation activities should be complemented with public surveys that shed light on the use of cultural amenities by the entire city population, including non-subsidized cultural amenities in the Netherlands, performing-arts venues use the Theater Analyse Systeem (Theatre Analysis System) as a benchmark, see [http://www.vscd.nl/dossiers/16/Theater_Analyse_Systeem_TAS/](http://www.vscd.nl/dossiers/16/Theater_Analyse_Systeem_TAS/) accessed on 28 July 2009.
facilities and amateur arts. Until now, reception research seems not to have been included in policy evaluation. As a result, this stage represents a major improvement in arts policy evaluation practices in the Netherlands.

Stage 4: Integrated policy evaluation
Key Question: do the combined efforts of cultural, sports, welfare, education and spatial planning policies yield changes in social capital and economic performance in our city?

The most elaborate stage in policy evaluation supplements the previous stages with information on the societal outcomes of cultural policy. These outcomes can only be measured on a higher level of aggregation and involving all influences on social and economic performance of a city. Therefore it is logical to use this stage to evaluate the joint efforts of cultural, sports, welfare, education and spatial planning policies. The evaluation will aid the city administration to align the efforts of these policies. Note that it is possible to conduct more focused research on the impact of aesthetic experience on social capital for specific groups who have been subject to specific cultural interventions, e.g., community arts projects. The only evaluation efforts in this stage are economic impact studies. However, none of these has been conducted with the right methodologies. This stage therefore represents a major improvement in arts-policy evaluation as well.

12.4. Summary

The model for cultural policy evaluation presented above is more elaborate than the current policy evaluation methods. According to this framework, evaluation can be considered more intrusive for artists. Although objections to more intrusive evaluation measures have been voiced (see the Introduction), the model can help both policy-makers and managers of cultural institutions in assessing how the arts function in society, and under which conditions, so that art’s functioning can be improved. This is exactly how Boorsma and Van Maanen perceive the usefulness of research into the occurrence of aesthetic experience. Such investigations can ‘provide important managerial information for arts organizations. It helps to evaluate whether artistic goals are reached, and it can produce insights to improve the conditions that facilitate the reception of their products in the future’ (Boorsma and Van Maanen, 2003a, p. 159), provided that the artistic integrity of the organization is not compromised; i.e., that the information is not used to adapt performances to consumers’ preferences (ibid., p. 184). The notion of art’s autonomy in society stipulates that audience and reception research is carried out by the arts institutions, giving them the first opportunity to interpret the data and to decide on the consequences for an institution’s strategies.

The second step is to report the data to the subsidizers. As Birnkraut (2008) suggests, the measures in the model for arts policy evaluation should not be used as cut-off rates that
immediately penalize underachieving institutions. Rather the evaluation techniques suggested can shed light on how the specific contribution of the (performing) arts to urban society can be realized and how its functioning as art can be strengthened through better primary (i.e. artistic) and secondary (i.e. management, financial and marketing) strategies of subsidized institutions. It is up to the city administration to use the information from these evaluations for public accountability by presenting the aggregated results.

Therefore it seems that, for evaluating (performing) arts policy - focusing on art’s specific contribution to society seriously as the model presented here does –, city administrators can do more than simply measure the output of (performing) arts institutions in terms of the amount of performances, attendees, and an assessment of the artistic quality of productions. However, evaluating cultural policy is a multi-layered activity that can be time and money-consuming. The evaluation model presented here supplements current arts policy evaluation methods in several ways:

(a) The occurrence of intrinsic values at personal level.

Information on the values created for audience members is rarely gathered. This means that the core of what cultural policy is about is almost never evaluated. This seems a surprising conclusion but it is consistent with the observation that cultural policy instruments depend on the institutional approach even although the legitimization of cultural policy is based upon the functional approach to art. In the institutional paradigm, the peer judgement of cultural values is crucial. Expert judgements on artistic quality form the core of subsidy allocation and also of cultural policy evaluations. As demonstrated in this research, this is a very limited approach to cultural policy evaluation.

(b) The occurrence of functions in the social domain

Although a link between aesthetic values at personal level and functioning in the social domain is frequently suggested in the policy documents, this link has never been researched in policy evaluation efforts. This is surprising because of the emphasis on personal development as a function of the arts and the implications of this development at societal level in the policy documents. However, the lack of a robust policy theory in this domain and the issue of causality can explain the current omission of evaluation efforts in this respect.

(c) Measuring organizational capacity

Secondary policy evaluation has been a recent trend in city administration, as can be witnessed in the evaluation efforts of Groningen and Rotterdam. Although many indicators can easily be identified and measured in this area, efforts to report on these issues on an aggregate level should be undertaken by city administrations.
12.4.2. Meeting the Challenge of evidence-based policy

In the Introduction to this research, New Public Management has been described as an ambiguous phenomenon. On the one hand, it is positive: who can be opposed to transparent, effective and efficient government? On the other hand, its basic assumptions are simplistic and require intricate social phenomena to be expressed in simple quantifiable indicators, while assuming straightforward cause-and-effect relationships behind such phenomena. The present research also suffers from this ambiguity. The framework for describing the functioning of the performing arts in society (graphically depicted in Figure 10.1) does cover the complexities of society, as it shows how the various values and functions are interdependent in a system of values and functions. But it still proposes an optimistic (or positivistic or modernistic) view of the performing art’s functioning in society. The framework should be ‘read’ from left to right suggesting a clear direction of cause-and-effect relations, even ignoring negative outcomes although not denying them. The discussion of measurement issues considerably qualifies the assumptions of the measurability of societal phenomena. It has been demonstrated that measurement is no easy task; it may even be very costly. But still, the main tenet behind the framework is that – apart from some omissions and with considerable mitigation of the assumed causality – measurement of the functioning of the performing arts in society is possible, providing that some of the evaluation tools needed will be expanded in the near future. Most notably, the questionnaire needed to assess the occurrence of aesthetic values at personal level should be developed on the basis of the now-available techniques. Furthermore, investigation should be carried out into whether or not participation in the performing arts yields specific types of social capital and how these specific types can be measured.

Collectively, the measurement instruments discussed in this chapter can assess the realization of values in society as a result of the aesthetic experience of the performing arts. Although an exact measurement of the ‘amount’ of values and functions realized is not possible, these measurement activities jointly ‘paint a picture’ of the functioning of the (performing) arts in society. Rather than having a thermometer to measure the performing arts’ functioning in society, these measurement activities provide a barometer to assess developments in the arts’ functioning in society.

The present research contributes to the three fundamental issues regarding NPM that were identified in the Introduction:

Can art be an agent of social change?

The framework for describing the functioning of performing arts in society demonstrates how art – or rather: aesthetic experience – can serve instrumental goals and thus can contribute to social change. Specifically, it indicates how the artistic nature of aesthetic experience contributes to this. However, the non-artistically aesthetic values and functions are intrinsic as well. Although they are not unique to aesthetic experience, non-artistically aesthetic values and functions are relevant in evaluating the impact of art in society. The
framework demonstrates that the present research falls squarely into the positive tradition regarding art’s impact on society.

**Can or should arts policy be instrumental?**

The present research does not provide a clear answer here, as this primarily is a political question. However, it does help in the debate on the evidence base of art’s contribution to instrumental goals, as the framework provides a sound theoretical basis for this instrumentality and it provides measurement methods that can be used to start building the empirical evidence base which is lacking at the moment.

**Can or should art’s functioning be measured?**

The answer here is Yes, though not without considerable effort on the part of city officials and managers of cultural institutions. The model presented here provides a framework to discuss coherently the impact of the arts. The model also presents limitations to measurement because:

- There is a diminishing causality when moving from left to right in the framework (Figure 10.1).
- Although the intrinsic functions at personal level can be measured (through a combination of quantitative and qualitative longitudinal reception research), the problem of timing of the evaluation research remains.\(^\text{23}\)
- Some of the measurement techniques necessary have not been fully developed. Thus it is questionable whether or not the model for arts-policy evaluation presents a clear enough vocabulary for officials and managers. However, it does provide a common system of thinking in evaluating the performance of institutions in relation to the goals of the public policy.

It should be noted that those who lean towards the modernistic side of the NPM debate may not be convinced by this model for arts-policy evaluation, as it does not indicate how to measure *the extent to which* the values that are relevant actually occur. Nor does it offer a yardstick or cut-off level to be able to say that the functioning is *sufficient* to legitimize arts policy. This, too, remains a political decision and therefore Gray (2007) and Belfiore (2004) are right in being wary of evidence-based policies. This is not particularly problematic for the Dutch situation, as the discussion of the performing-arts policy shows that the policies are primarily legitimized on the basis of the intrinsic values and functions at personal level, and that economic and social impact of the performing arts has a side-effect status.\(^\text{24}\) The discussion also showed that the three tendencies Gray (2007) postulates on the basis of his commodification thesis have not fully been materialized in Dutch arts policy during the period researched. The policies are still rather general and not focused on specific groups

---

\(^{23}\) When research delivers no evidence for the occurrence of intrinsic functions at personal level, it cannot be excluded that they might occur at some point in the future.

\(^{24}\) With the exception of the early stages of arts policies in cities, when policies tended to be legitimized by extrinsic arguments.
and any specific impact. Supply subsidy is still the dominant policy instrument. This leads to the question as to whether or not the arguments against instrumentalism have been voiced in terms that are too general. The anti-instrumental rhetoric also simply claims that art’s impact is intangible and does not recognize that art, when functioning as art, can indeed contribute to wider societal goals.

Basically, this research shows how, through a functional (or pragmatic) perspective on art – or rather, aesthetic experience –, New Public Management measurement requirements can be met in a way that it respects the specific nature of art. It shows that, by their very nature, arts policies are instrumental because all intrinsic values associated with aesthetic experience ultimately contribute to (not lead to) non-aesthetic means, precisely as a result of the aesthetic values and functions of the experience. This is a fundamental ambiguity of arts policy that anyone responsible – politicians and city officials, artists and managers of cultural institutions alike – will have to deal with. Especially in the brave new evidence-based policy world we live in today.

However, one should question whether or not the recent policy attention to folk culture and national history as a means to orient the nation to the Dutch identity (whatever that may be) – which has been accompanied in national politics by a discussion on the new National History Museum that will have shaken the believers in art and culture’s autonomy to their very core – and the inclusion of culture as an instrument in the regeneration of the ‘Vogelaarwijken’ (40 city boroughs across the nation targeted for regeneration) herald a fundamental shift towards stricter specification of targeted groups and impacts.