WHO TELLS FORTUNES TO NONPROFIT PROFESSIONAL THEATRE PRODUCERS? STATE INTERVENTION AND ITS STRINGS ATTACHED

Paper submitted for the XII International Conference on Arts & Cultural Management, AIMAC 2013
26 – 26 June 2013, Bogota, Colombia

-In case of interest in citation, please contact the author for an updated version of this paper-

ELLEN LOOTS,
ellen.loots@ua.ac.be
PhD-student, University of Antwerp, Belgium

Ellen Loots is a PhD-student at the Antwerp Centre of Evolutionary Demography (ACED) and the Competence Centre Management, Culture and Policy of the University of Antwerp, Belgium. She studies the performance of organizations in arts, with a specific focus on mediating and gatekeeping processes.

ABSTRACT

The performance of organizations is a major yet complex issue in management and organization studies. In the arts that are situated at the non-profit end of art production, implying the involvement of several stakeholders such as the state and professional groups, performance definitions and measures are equivocal. The present paper investigates the evaluation process that a patronage-giving state has developed in order to proficiently judge the worthiness of applicants to receive public support. By combining content analysis and statistical analysis we elicit a discrepancy between the verbal evaluation (the written reports) and the underlying decision-making process of panelists. *Artistic quality* is the major determinant of subsidy allocations to the detriment of *social concerns*. The absence of more direct effects in our analyses is striking and the conclusion that only a fraction of the imposed criteria matters is conspicuous. We probe some implications for the performance of theatre producers, which need to balance between mission and survival.

Performance, theatre producers, evaluations, allocation decisions, expert panels

INTRODUCTION

The performance of organizations has been a major yet complex issue in management and organization studies. Back in 1986 Venkatraman and Ramanujam (1986: 801) already addressed the complexity of defining performance: "Although the importance of the performance concept (and the broader area, organizational effectiveness) is widely recognized (...), the treatment of performance in research settings is perhaps one of the thorniest issues confronting the academic researcher today. With the volume of literature on this topic continually increasing, there appears to be little hope of reaching any agreement on basic terminology and definitions." In the last decades some dominant performance indicators have become salient in the canonical literature on the assessing and explaining of organizational performance. In the majority of industries goals have a major financial part, because profit maximization is the ultimate objective of most firms. This is realized and corroborated through instances such as optimal production, sales maximization and customer service. But since organizations have multiple and frequently conflicting goals, the definition of performance remains a challenge for researchers in any field of management (Chow, Heaver, & Henriksson, 1994).

In arts, especially in the arts that are situated at the non-profit end of art production, performance definitions and measures are equivocal. This can be explained partially by the involvement of several stakeholders. Applicable to most art forms is the idea that artistic creativity is not as much an act of the individual genius as it is the outcome of a collaborative effort by many, as expressed by sociologists with Howard Becker a forerunner in the 1970's (Becker, 1974). Several have studied the critical roles of a range of actors that are complementary to the artist in the creating and displaying of works of art (Caves, 2000; Hirsch, 2000; Voss, Cable, & Voss, 2000; White & White, 1993 (1965)). It is easy to imagine that the interests and values of these actors not necessarily align. As such, the field of arts or cultural production (Bourdieu, 1983; Peterson & Anand, 2004) can be comprehended as one at "a juncture of different societal logics" (Goodrick & Reay, 2011: 406). The field of cultural production has become affected by logics of the market (consumer demand) as well as by that of the professions (Greenwood, Diaz, Li, & Lorente, 2010; Oakes, Townley, & Cooper, 1998), that in arts do not only include artists, but likewise arts managers, distributing companies, critics and other experts. Besides that, also the state and public officials are strongly involved, especially in the supporting of art that is not for profit and may have a public good character. All parties would want an audience to meet the art produced, for sure; but the endogenous dynamics navigating each, such as interests, values, power dependencies, and capacity for action (as categorized by Greenwood and Hinings (1996)) do not necessarily align (Krebs & Pommerehne, 1995). If performance may be defined as the extent to which predetermined goals are achieved (Chow et al., 1994), we can imagine that the involvement of several stakeholders – including taxpayers and their political representatives – may lead to a multifaceted possibly conflicting bundle of performance goals that art producers are confronted with. These goals beyond the control of the producers are outspoken or candid, and they may serve different functions, which extend far beyond the production of high-quality arts. Performance evaluations are steered by different motivations and may take place for different reasons, including the monitoring, evaluating, affecting and inferring of the behavior of arts organizations (Schuster, 1997).

What are the implications for the producers of arts? On the one hand, arts organizations can be understood as commodity producers, that generally attempt to convert all actions into the buying and selling of commodities, or converting an audience to their artistic offering (Friedland & Alford, 1991). On the other hand, relying upon revenue grants for their operations, these organizations to some extent resemble public sector organizations that must be organized in accordance with state logics, as one of imposed efficiency (a business-like logic of cost-effective operations) as well as one of meeting particular needs of citizens (Reay & Hinings, 2005, 2009). In democratic states, cultural policy is aimed at the provision of art and contended to be contributing to the wellbeing of citizens (Feder & Katz-Gerro, 2012). Policymakers seek to justify the distribution of public means, for which they develop a rationale and a conditional rule-set. So grants to organizations have strings attached in that they stipulate how an organization is expected to behave, at least on some dimensions. Thereto, most governments foresee an assessment procedure, that is partially business-based, partially oriented on the artistic output. Public donors have been said to define the conditions for public support in terms of variables that are easily observed, such as capacity utilization (Krebs & Pommerehne, 1995). Nonetheless, these performance evaluations by public donors, both ex ante and ex post grant allocations, regularly intend to also assess the intrinsic quality of an organization's merchandise (Schuster, 1997) and outcomes rather than outputs (Pignataro, 2011).

In the present paper, we study the granting decision process in two democratic regimes in continental Europe: the Netherlands and the Dutch-speaking region of Belgium (Flanders). More specifically, we study the relationship between the conditions to theatre producers as put forward by a state, and the evaluation of these aspects as assessed by expert panels that are prearranged to evaluate organizations against a set of formal criteria. One of the dimensions under scope is artistic quality, the major concern of many of the grant-applicants because it is their reason of existence.

The remainder of the present paper is as follows. After clarifying the method that is used for this analyses, we expose the context of our inquiry. We introduce the policy requirements toward subsidy-seeking arts producers as criteria of an evaluation process. The subsequent steps of our analysis consist of counts, correlation analyses and regression analyses in order to investigate respectively the transparency of the evaluations, the requirements to art producers to be legitimate actors and the determinants of reception into state sustenance. In a conclusive section we probe a number of explanations for the processes at hand and we induce implications for arts producers.

METHOD

Abduction

For this inquiry we applied what has been referred to as an ***abductive logic***. Charles Sanders Peirce suggested the logical category of abduction in research. It provides criteria for scientific discoveries distinct from ***criteria of proof***. Where deduction is a process of checking theories, and induction one of checking hypotheses, abduction can be conceived of as a method for the discovery of hypotheses ('on probation') or propositions, the latter being abductive explanations of surprising facts (Levin-Rozalis, 2000). The process of abduction starts with the observation of a new or surprising phenomenon. It is then necessary to add interpretations to the facts, or assumptions about possible interpretations (propositions), in order to adapt them to situations that are different from those in which they have been observed. Especially in social studies that have a main purpose of evaluation and in which attempts are made to give effective explanations for field-dependent phenomena, an abductive method is convenient.

Content analysis

We use the written judgments of the evaluation committees as a major source of information, from which theorizing can emerge. Thereto we apply content analysis (Kassarjian, 1977; Kolbe & Burnett, 1991; Krippendorff, 1980). Content analysis has been defined as "***a research technique for making replicable and valid inferences from data to their context***" (Krippendorff, 1980: 21). A close-up upon textual information is salient for this research for at least two reasons. First, we expect that in our context a specific vocabulary is used to evaluate art producers and their products. As in science, we expect instances such as 'original', 'innovative', 'important' or 'significant' to express valuation (Guetzkow et al., 2004). Second, the specific type of this text appeals for a profound approach, given its hermeneutical status. It is a written judgment that expresses the reasons for why one should (or not) be supported by the state, guided by a bundle of predetermined criteria. It is a judgment of field members on other field members, who need be able to look into each other's eyes afterward. So an in-depth reading of these texts is necessary because we expect these judgment reports to evaluate many aspects on different levels of interpretation.

Content analysis implies a quantification of text excerpts. Thereto a coding scheme is to be developed, for which a first step is the defining of content categories. In our study these categories are the criteria set by the government. Since they are short in commentary and have no descriptive clarification[1], we had to produce a set of reliable and valid indicators that expressed the distinguished content and dimensions of each criterion. Also to determine is the basis unit of text to be selected and classified, which could be paragraphs, thematic units, sentences or words (Strijbos, Martens, Prins, & Jochems, 2006). We decided to use thematic units on the level of the sentence or block of sentences. The selecting of the content and attributing it to a category was an iterative, recursive and reflexive process. Independently two researchers coded the texts making use of Nvivo qualitative research software. During the analysis 'dictionaries' were developed. These consisted of lists of words and phrases that were associated with the categories of the content we looked for. After the coding the results were pooled and the final retentions were discussed and agreed upon.
The coded texts led to numeric variables. For each organization two dummy variables were developed. A first variable v1 indicates the presence (1) or absence (0) of a topic on the level of the organization. It expresses whether the evaluation panel assessed an organization to a specific criterion, or if it neglected to do so. A second variable v2 is a valued judgment: we assigned the value of one when the panel was positive about an organization on a singular dimension and the value of zero when the panel was negative. The nature of the data allowed us to code 'artistic quality' into a more fine-grained scale of 5 values. The Average Deviation Index by Burke & Dunlap (2002) measures the dispersion of responses around the mean; we obtained a value of $AD_{M(J)}$ $AD_{M(J)}$ = 0.43, which can be considered indicative of acceptable inter-rater agreement.[2] If not similar, the scores of the two raters were averaged.

Data

1 Except for the artistic quality variable in the Netherlands

2 An AD_{UL} AD_{UL} value (UL = upper limit) for 5-point items scales is 0.8 (the cutoff value of 0.8 is statistically significant at the .05 level for all Likert-type scales with five or more response options, as long as the sample size N equals or exceeds 13) (Burke & Dunlap, 2002).

We study the evaluations of 84 Dutch and 57 Flemish theatre organizations that have applied for a subsidy.[3] The money is intended to contribute to the operations by filling-up the yearly deficit, which can mount up to a share of 70 to 80% of turnover. Kindled by the New Public Management movement the public sectors in the Netherlands and Belgium have introduced performance management systems and open communication on policy decisions, budget allocations and outcomes. In both countries the application and assessment process under scope takes place every four years. Theatre companies and theatre houses can submit an application file that is being assessed on its financial and administrative plans, and on its artistic plan. For the latter expert panels are invoked. These panels consist of individuals that are attributed an institutional responsibility to make judgments on the artistic performance in arts. Panel members originate from within the country and are seen as experts in the domain. Members know the criteria to use, but they are generally not instructed on how to interpret the specific criteria or on how to ponder each one. They are only accountable to other panel members and no rules on how to evaluate are written down. Panel members arrive at a consensus in a meeting.
The policy program goals are set in legislation and manuals, which also inform the applicant pool. The criteria against which they are assessed ground in these policy documents. In Flanders the Arts Decree stipulates the political priorities related to performing and visual art forms, and the conditions under which organizations can apply for government support in the form of direct grants. A handful criteria is related to structural aspects of the applicant, such as legal personality, operating base and minimum staffing (Arts Decree, subdivision II, article 7, §1, 1 to 11). Additionally some minimum standards are being postulated, such as a minimum number of activities, a minimum share of market income, and a minimum share of turnover to be spent on wages (Arts Decree, subdivision II, article 7, §2). If compliant to these formal requests, an applicant needs to give proof of acting in accordance with a bundle of major substantive criteria (Table 1). In the Netherlands a corpse of funds execute a major part of national cultural policy by granting subsidies to organizations and individuals. The Dutch Funds for the Performing Arts has the task to realize a diverse, qualitatively high and well-dispersed supply, which it in the time window under scope developed under the motto 'more for less', implying a stringent selection on the one hand, and an increase of the average subsidy on the other hand. The Dutch regulation stipulates that applications are being judged upon five main aspects, of which some of the content is elaborated. In general, the judgment on the quality is based upon *'the quality of the previous work and/or activities, the qualitative development of the work and/or activities, and the plan and budget.'* The first quality-criterion refers to the artistic quality. Three central notions to the evaluation of this artistic quality have been put forward and made explicit: craftsmanship, originality and expressive power. Opposed to artistic quality, criteria 2 to 6 are being labeled the 'functional quality'. This relates to the way in which productions and performances function within the Dutch society.
Table 1 depicts the dictionaries of the criteria as they have been used to do the content analysis of the judgments. The Flemish criteria were hardly elaborated; the interpretation came into being during the analysis. The Dutch content is largely based upon the descriptions provided within regulatory documents. Four major performance categories underlying the criteria can be recognized: artistic quality, a social dimension, efficiency and strategic premises.

3 Belgium is subdivided into three language-based communities that organized some dimensions of policy separately, such as culture and arts. We focus on Flanders, which is the Dutch-speaking northern region, neighboring to the Netherlands. We study the previous term, based on decisions of 2008-2009.

Table 1: evaluation criteria for theatre organizations in Flanders and the Netherlands

Criteria	Belgium (Flanders)	The Netherlands	PERFORMANCE DIMENSIONS
ARTISTIC QUALITY	Creations that are intriguing, innovative, strong, emotional, beautiful, relevant, original; honest and authentic theatre; important, inventive, talented, devoted creators; a theatre language that is expressive, unique, original, important, strong; devotion to the craft of theatre; the dramaturgy is strong/poor; texts are original, made accessible, communicative... (3)	craftsmanship, skills of team members, originality, artistic signature, vision of the artistic leader, distinguished, expressive power, challenge imagination, the creation of meaningful performances (1)	Quality and reputation
PROFILING & POSITION	sharp/strong profile, unique position, special, coherent, (no longer) relevant or unique, remarkable, original, obstinate, having played a (pioneering) role, having been part of..., one of the representatives of..., distinguishing... (1)	different, unique (2)	
(INTER)NATIONAL REPUTATION	Spreading, dispersal, dissemination, international operations, abroad, (limited) visibility, presence/absence (in specific countries, in national theatres...), touring (within a country, around national theatres...), travelling to, (inter)national reputation, increase of attraction, disposing of a network of European partners, mapping internationally (also 1), crossing regional/national borders (4)	outreach, dispersal, distribution (3)	
REGIONAL ADDED VALUE	regional meaning/significance, regional recognition, radiance, embeddedness, anchoring (10)	link to contribution to the distribution (3)	Social value and market
AUDIENCE OUTREACH	target groups/audience, arts education, public relations, audience recruitment/growth (7)	Audience development, education, marketing (5)	
DIVERSITY AND INTERCULTURALITY	New Belgians, multicolored and multilingual environment, non-western cultures, diverse backgrounds (13)	cultural diversity (intercultural supply and program) (2)	
SOCIAL VALUE	social-artistic activities/dynamics, social importance, socially critical and political themes/language/voice, social engagement (11)	-	
COLLABORATION AND NETWORKING	collaborations, network, alliances, co-productions, synergies (5)	-	Efficiency
CHAIN IDEA	-	the creation, production, programming and the audience development, exemplified by collaborations (4)	
CULTURAL ENTREPRENEURSHIP & MANAGEMENT	-	market income, playlist, attitude (5)	
LONG-TERM VISION	plans, (lack of) vision, policy (2)	-	Strategy and planning
REALISM	we do not find explicit synonyms, but do find combinations of words that reflect this instance (9)	-	
FEASIBILITY	the plan is (not) feasible (6)	-	
() between brackets is the number of the original criterion			

RESULTS

The comprehensiveness of the evaluations
Theoretically, panels can be expected to assess organizations on each single criterion imposed by the government, but in practice evaluations might not happen as formal and tight. This section studies the occurrence of criteria that are being put forward by the government (variable v1). It likewise partitions the criteria in relevant components that emerged from the texts. Table 2 displays the incidence of the various criteria within the judgments.

			Belgium (n=57)		freq	The Netherlands (n=84)		freq
Criteria	dimensions		Presence N	%	(rank)	Presence N	%	(rank)
Artistic quality			**57**	**100%**		**84**	**100%**	
Profile and position	profile or position		**47**	**81%**	***5***	**84**	**100%**	***1***
	profile		29	50%		17	20%	
	position		38	66%		25	30%	
		unique	21	36%		12	14%	
		distinctive				16	19%	
		complementory				26	31%	
		meaningful				4	5%	
		significant	30	52%				
		location	10	17%			0%	
		important	6	10%		8	10%	
		strong	5	9%		1	1%	
		remarkable	1	2%			0%	
		key player	3	5%		3	4%	
		recognizable				3	4%	
		visible	6	10%			0%	
		exemplary				2	2%	
		pioneering (first)	3	5%		5	6%	
		innovative				2	2%	
		mature				1	1%	
(inter)national reputation			**53**	**91%**	***1***	**36**	**43%**	***6***
	national		40	69%		26	31%	
	international		40	69%		22	26%	
regional added value			**28**	**48%**	***8***			
	region		16	28%				
	City		12	21%				
audience range			**49**	**84%**	***4***	**54**	**64%**	***2***
	children		11	19%		9	11%	
	youth		18	31%		15	18%	
	traditionally high		4	7%		2	2%	
	traditionally low		12	21%		3	4%	
	new		11	19%		21	25%	
	social groups		19	33%		4	5%	
	divers		29	50%		9	11%	
	large					22	26%	
interculturality & diversity			**40**	**69%**	***6***	**22**	**26%**	***9***
social value			**38**	**66%**	***7***	**24**	**29%**	***8***
	content		22	38%		18	21%	
	outreach		23	40%		10	12%	
collaboration and networking			**52**	**90%**	***2***	**51**	**61%**	***4***
	to economize		16	28%				
	to create staging opportunitie		41	71%				
	to develop (coproductions)		45	78%				
chain idea						**54**	**64%**	***3(2)***
cultural entrepreneurship and business						**42**	**50%**	***5***
long term vision	vision		**22**	**38%**	***9***	**26**	**31%**	***7***
	plan		35	60%		18	21%	
realism			**11**	**19%**	***10***	**6**	**7%**	***10***
feasibility			**50**	**86%**	***3***			

Table 2: the occurrence of the criteria in the evaluations in Flanders and the Netherlands
Each organization is in each of both countries under scope being evaluated on its *artistic quality*, which may not come as a surprise. Although both governments put forward a dozen criteria to which applications have to conform, except from artistic quality not one criterion is used to assess each single case upon. *Reputation*, which is connected to past performance, is the second most heavyweight criterion in terms of occurrence in Flanders. This could imply that newcomers might have difficulties entering the subsidy system. Whether or not a plan is *feasible* is being reported in 85% of the cases. This may be a decisive criterion. Yet, we could question the salience of this criterion, because the panel only evaluates an artistic plan, and not its financial counterpart. Other criteria that are frequently invoked in Flanders are *profile and position*, and *audience orientation*. *Position* implies many distinct dimensions such as a geographical, unique, prominent, pioneering, or visible position. Also in the Netherlands some distinct dimensions of *profile and position* come to the fore. A subset of organizations is being assessed on its uniqueness or distinctiveness. In the Netherlands the expression 'complementary' is being used; in Flanders the notions 'added value' and 'significance'. In addition to that, there appear to be some strong actors (important, remarkable key players), and some exemplary (or innovative or pioneering) actors. Two criteria remain underrepresented in the Flemish evaluations: *long term vision* and *realism of a growth path*. It appears as if features that are related to the content of the grant application, often projecting a future plan, do not prevail to an organization's reputation, which is mainly built upon past performance. This finding is in line with Gorman (2007) and Guetzkow et al. (2004) that illuminated similar tendencies within evaluations in science, referred to as the 'Oppenheim effect'.[4] At the tail of the list of criteria in Flanders and not an official Dutch criterion, is *vision*. A close reading of the texts suggests that in both countries a lack of vision may be a sufficient condition for rejection.
In sum, we see that in Belgium there is more consistency in the application of criteria, even though they are manifold, and open-ended. In the Netherlands where fewer criteria are to assess, less perseverance in their application can be noticed. Profile and position are austerely being considered in these assessments, giving the impression that panels perceive as its major task to compose an artistically diverse landscape of complementary organizations, in a way that scenarists would design a set.

Features of legitimate actors

The correlations between the variables (v2) will clarify two things. Correlations between criteria and judgments will first reveal which criteria appear to be counting in the decision making process. Second, significant correlations among criteria will disclose which criteria relate and might co-occur within the assessments. If legitimacy can be understood as the outcome of being evaluated as doing the things appropriate to do, these correlation coefficients give insight in the characteristics and activities of legitimate actors in the theatre fields in Flanders and the Netherlands.[5]
The correlation between *artistic quality* and the judgment is highly significant. This is the case in both countries, with the Netherlands (0.887) outperforming Flanders (0.778) on this issue. This finding implies that the assessments are a clear reflection of the artistic quality of an organization, as it is perceived by a group of experts entitled by the state to assess drama organizations on their performance. Yet the correlation is not ample; some unexplained variance of 10% to 20% remains. This suggests that some judgments are advanced despite of an artistic evaluation in one or the other direction and that other (latent) factors may be at stake.
A remarkable difference between the two countries relates to *audience orientation*. It is significantly though moderately correlated to the judgment in the Netherlands, while in Flanders any relationship lacks. The lack of relationship could indicate that a large or specific audience base does not have any account in the subsidization of theatre in Flanders. Another criterion that both states apply is *diversity*. The diversity focus in Flanders is somewhat ambiguous in its relation to the outcome, whereas in the Netherlands, albeit less taken into account, it is significantly correlated to the arbiter's judgment. The criterion of *feasibility* is correlated to the judgment, even on a fairly high level (0.576). Two other criteria that might reflect upon future undertakings, *vision* and *realism*, are more contested in relation to the evaluation outcomes (low values and moderate significance), as is *collaboration*. Collaboration shows a negative sign, suggesting that collaborations would be counterproductive in order to receive a positive assessment. This, however, is not in line with political inclinations by the Flemish authorities that instigate on collaborations in order to keep down the expenditure to production. Likewise in the Netherlands a claim for effectiveness has been converted into evaluation standards: for an organization to be

4 The Oppenheim effect in the peer reviewing process by scientific journals describes the phenomenon that the author and not the manuscript itself is the criterion in determining quality to the submission and is the base for the editorial decision (Gorman, 2007).

5 Many scholars have written about legitimacy and its different forms, antecedents and consequences. We here conceive of legitimacy as it was defined by Suchman (1995), as moral legitimacy, which is a form of legitimacy based on the normative approval by an expert audience (Suchman, 1995).

subsidized, it should express *cultural entrepreneurship* and inscribe within the Dutch "*chain idea*", a principle that tries to organize the field in an efficient and cost-effective way. The latter is only moderately correlated with the judgment, as opposed to *cultural entrepreneurship*. At last, two substantive priorities of the concurrent policy maker in Flanders give the impression of not being considered by evaluators. An actor's *social* and *regional value* can be considered to be political values that do not align with an expert committee's priorities, be it by a lack of interest or a lack of knowledge or another latent motivation.

Looking at the pairwise correlations between the criteria, we see that in both countries three criteria significantly correlate: *artistic quality, position* and *reputation*. This is not an unexpected finding, since artistic quality could be an imperative determinant of one's reputation and the position one uptakes. In Flanders, an additional criterion adds to this triangle: feasibility. The relation is less straightforward. An evaluating body could obviously trust upon previous achievements by an organization, to gauge whether it will be able to deliver what it projects. But, as advocated earlier, feasibility could be a key determinant, functioning as a barrier to entry or exit from the subsidy system. In addition to scoping upon the presence of correlations, it is also salient to look at the absence of correlations. In Flanders, the criterion of *regional value* does not relate to any other variable. This might give proof of a discrepancy between a (mere) regional value, and the worthiness of grant allocation: grants may be privileged to organizations that attain such a level of overall quality that they perform beyond regional or local latitude. Also criteria that relate to *audience orientation*, *social orientation*, and *diversity*, all recent political priorities do not coincide with values that are typically associated with the high arts: *quality*, international appeal or *reputation*, and *position*. These findings suggest that the value systems or norms of expert panels and grant allocators do not align. The nature of the discrepancy recalls the type of division that Pierre Bourdieu pointed to several decades ago (Bourdieu, 1979).

Table 3: correlation table of the criteria in the evaluations in Flanders

Means, standard deviations and Spearman's Rho Correlations among all variables

	Mean	S.d.	1	2	3	4	5	6	7	8	9	10	11	12	13	14
independent variables																
1 Artistic quality	3,31	1,158	1,000													
2 Position	,66	,479	,572**	1,000												
3 Reputation	,74	,442	,616**	,566**	1,000											
4 Audience	,84	,365	,125	,090	,291*	1,000										
5 Diversity	,52	,504	,155	,243	,217	,062	1,000									
6 Vision	,31	,467	,252	,330*	,311*	,082	,201	1,000								
7 Feasibility	,69	,467	,528**	,533**	,455**	,021	,396**	,208	1,000							
8 Realism	,19	,395	,206	,166	,185	,086	-,061	,056	,039	1,000						
9 Collaboration	,69	,467	-,211	-,173	-,056	,021	,023	-,033	-,208	-,341**	1,000					
10 Regional value	,47	,503	-,008	,023	-,238	-,173	,141	,121	,178	-,099	-,046	1,000				
11 Social	,57	,500	,021	,174	,122	,204	,274*	-,093	,093	-,023	-,057	,114	1,000			
dependent variables																
12 Judgment (before response)	3,53	1,341	,779**	,662**	,543**	,144	,307*	,268*	,652**	,250	-,290*	,089	,109	1,000		
13 Judgment (after response)	3,67	1,316	,778**	,679**	,531**	,115	,232	,295*	,576**	,331*	-,281*	,046	,106	,970**	1,000	
14 Judgment (after, dummy)	,78	,421	,644**	,741**	,438**	,112	,308*	,361**	,622**	,155	-,271*	,170	,200	,735**	,756**	1,000

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

Table 4: correlation table of the criteria in the evaluations in the Netherlands

Means, standard deviations and Spearman's Rho Correlations among all variables

Variables	Mean	S.d.	1	2	3	4	5	6	7	8
independent variables										
1 Artistic quality	2,51	,844	1,000							
2 Position	,33	,473	,528**	1,000						
3 Reputation	,43	,498	,611**	,340**	1,000					
4 Audience	,66	,477	,356**	,121	,309**	1,000				
5 Diversity	,27	,446	,341**	,220*	,257*	,262*	1,000			
6 Chain idea	,63	,485	,102	,209	,195	,094	,117	1,000		
7 Cultural entrepreneurship	,35	,481	,406**	,350**	,290**	,264*	-,045	,244*	1,000	
dependent variable										
8 Judgment	,50	,503	,887**	,597**	,715**	,463**	,385**	,253*	,485**	1,000

**. Correlation is significant at the 0.01 level (2-tailed).
*. Correlation is significant at the 0.05 level (2-tailed).

The determinants of state sustenance

A final statistical estimation will elicit whether evaluation panels actually invoke the criteria that are advanced by a state. In theory this should be the case, because the funding bodies present them to applicant-organizations as a non-exclusive list without any weights.[6] We would expect significant coefficients of the same magnitude for all the variables in the model. The correlations however suggest that only a subset of criteria affects the panels' decisions. Still, the written evaluations give the impression that the panels assess organizations on several dimensions. Yet, there could be a discrepancy between the rhetorical argumentation of the panel, and its decisive behavior.

In the Netherlands, the evaluation panel expresses its evaluation straightforward, as a 'yes' or a 'no' to subsidy allocation. For the analysis of the Dutch data we thus develop a *multiple logistic regression* model that describes the probability of being evaluated positively by a panel of experts as a function of the criteria advanced by policy makers (Cohen, Cohen, West, & Aiken, 2003). A bias-reduction method was needed to overcome a data separation problem (Firth, 1993), due to the dominance of one variable (artistic quality). Running this corrective method in R by the package logistf (Ploner, Dunkler, Southworth, et al., 2010),[7] we find that from all the variables only the effect of *artistic quality* is significant up to the 0.001-level; the effect of *reputation* is just below the threshold of 0.05, also in a second more parsimonious model; no other variable is significant (table 5).

	Model 1	Model 2
Criteria		
Constant	-7,920***	-11,970***
Artistic Quality	8,740***	24,779***
Position	1,875	
Reputation	2,393*	3,197*
Chain Idea	1,183	
Audience Orientation	1,78	
Diversity	0,435	
Cultural Entrepreneurship	-0,612	
Test statistics		
Overall model evaluation		
Likelihood ratio test	90,749****df7	97,632*** df2
Wald test	22,091**df7	7,842** df2
*=p<,05; **=p<,01; ***=p<,001		

Table 5: regression table (the Netherlands)

In Flanders, the evaluation panel uses a scale for assessing the eligibility for subsidies, ranging from 1 to 5. We apply the SPSS PLUM procedure that calculates a polytomous logistic regression model for ordinal dependent variables by making use of a cumulative logit model. Three variables appear to be significant for predicting the outcome: *artistic quality, position* and *realism*. A second parsimonious model only including these three variables exposes that the significance of *realism* dissolves. Compared to this reduced model, the full model with all eleven criteria explains relatively little additional value, adding only 3% to the pseudo-R^2 statistic.

6 In the Netherlands 'artistic quality' is advanced as the major criterion.

7 For a clear understanding of the phenomenon and the proposed solution we recommend the interested and mathematically inclined reader to address the original contribution of Firth (2003) and the methodological advances by Heinze and coauthors (e.g. Heinze & Schemper, 2002).

Table 6: regression table (Flanders)

	Model 1	Model 2
Criteria		
Threshold (judgment = 1)	-3,963	-1,695
Threshold (judgment = 2)	0,618	1,588
Threshold (judgment = 3)	2,319	3,03
Threshold (judgment = 4)	6,149	6,852
Artistic Quality	11,669***	10,680***
Position	-2,373**	-2,406**
Reputation	1,166	
Regional Importance	-0,485	
Audience Orientation	-0,327	
Diversity	-0,367	
Social Realm	-0,558	
Realism	-1,796*	-1,39
Feasibility	-0,918	
Vision	-0,797	
Collaboration	0,313	
Test statistics		
Overall model evaluation		
Chi²	83,561***df11	76,600***df3
Pearson	108,326 df 209 (p=1,000)	1,688 df 65 (p=0,885)
pseudo R²		
Cox and Snell	0,763	0,733
Nagelkerke	0,809	0,777
McFadden	0,503	0,461
*=p<,05; **=p<,01; ***=p<,001		

DISCUSSION AND CONCLUSION

The evaluation and measurement of the performance of arts organizations, which are subject to accountability toward several stakeholders, is a task fraught with difficulties (Hadida, 2013 (forthcoming); Turbide & Laurin, 2009). Evaluation criteria of different parties do not necessary align and may be tacit. The present paper investigates the evaluation process that a patronage-giving state has developed in order to proficiently judge the worthiness of applicants to receive public support. Our objective has been twofold. First, we explored artistic performance, while conceiving of it as a multifaceted concept that in a field as that of the arts may be influenced, maybe even piloted, by several actors external to arts producing organizations (Becker, 1984). Part of our focus was on the relation between *artistic performance* and *artistic quality*. Second, and more specific, we sought to clarify the relationship between performance requirements put forward by a state and the assessments by expert panels that are prearranged to evaluate organizations against formal criteria. This procedure can be considered as an instance where two distinct logics congregate: a better understanding of the underlying mechanisms could give insights in *how* a state logic and a professional logic are able to co-exist.

The application of a mixed-method approach in an abductive line of reasoning appeared prolific to our objectives. By combining content analysis and statistical analysis we elicit a discrepancy between the verbal evaluation (the written reports) and the underlying decision-making process of panelists. Our analyses show that a report containing an assessment of an actor against distinct criteria does not imply that these criteria matter to the final decision of the assessors. The redundancy of criteria that reflect social objectives has been illuminated in both Flanders and the Netherlands. A state may expect arts organizations to break down barriers in society and to concern with access maximization and social inclusion. However, for expert panels 'artistic quality' is a major determinant in the performance assessments, to the detriment of social concerns. In the Netherlands the 'artistic quality' variable is even a predictor of such magnitude that it produces a problem of *quasi complete data-separation*. The efforts needed to elaborate a convenient model for the Dutch data – the *Firth logistic regression model* (Heinze, 2002) – reasonably suggest the lack of a logic and consistent pattern underlying the data.

In the cases under scope ***artistic quality*** is the major determinant of subsidy allocations. Additionally ***reputation*** and ***position*** play a role. Both criteria have limitations. First, ***reputation*** is based solely on the perception of panelists. In order to avoid perception biases due to limited cognition, or even favoritism, more objective reputation indicators could be used. Second, ***position*** gave proof of being a multi-facetted concept. Panelists evaluate position haphazardly with regard to products and geographical markets; yet, applicants may have a unique position, due to either their distinctive products, either their unique presence in an otherwise deficient region. We advance that peer panels may be capable of determining the first dimension (originality), but that they lack the know-how of judging the second (dispersal). The country and region under scope are relatively small in surface; and our dataset of drama producing organizations combines distinct profiles, with larger municipal theatres on one side of the spectrum, and smaller touring companies on the other side. Several elements regarding the nature of the demand side of the market (as household preferences, customary travel distances and the supply in neighboring areas) should be acquainted with in order to gauge the complementarity of the cultural supply. The proficiency of 'expert panels' to judge about the complementarity or redundancy of distinct organizations in this or any other context should be questioned. Further, it is curious that these panels invoke the complementarity of the artistic supply as a decisive criterion. We suggest that the reason for doing so may be found within the scarcity of resources to be distributed. Our analyses are based on data from the years 2008-2009, in which the global economic crisis reached its apex: budget-cuts were warranted in all domains of public expenditure, and non-prosperous activities may have appeared especially vulnerable. We hint that the spontaneous yet strong pronunciation of the complementarity of cultural suppliers by these panels could have been a deliberate tactic aimed at the redeeming of the overall support to art. As if the experts uncover as peers.

Our investigation further accentuates a paradox in the evaluation process, which has been illuminated elsewhere: a plan is supposed to be decisive for grant allocation, but past performance reflected in instances such as artistic quality, position and reputation overrule future intents (Gorman, 2007). An undesirable consequence could be that incumbents are privileged, and thus that innovations will be hampered. We probe that past performance is an unavoidable yet strong determinant of decisions and selections in evaluation processes that primordially rely on the expertise of peer-panels.

The importance of the present findings lies primarily in the framing of an evaluation process that is developed by one party – the state –, yet executed by a third party – an expert panel –. The absence of more direct effects in our analyses is striking and the conclusion that only a fraction of the imposed criteria matters is conspicuous. Policy makers do not assume panels to assess different organizations with distinct measures and weights. Despite the existence of overt evaluation criteria, the performance evaluation in se is not transparently based on them. Of course, the present study is limited because it empirically investigates a particular evaluation process of one industry in only two countries. Still, we can advance some tentative explanations (hypotheses on probation).

The concept of institutional logics allows to address issues concerning different models of organizational functioning and accountability more clearly (Friedland & Alford, 1991). Studying the conditions that are attached to the granting of public funds and therefore focusing on the evaluations that are exerted by panels of experts, elicited the involvement in art consecration processes of two major actors - the state and the professions (experts) - that have been identified as the primary shapers of institutional forms (Dimaggio & Powell, 1983). The logics that these powerful actors attain may be distinct, but also able to co-exist (Goodrick & Reay, 2011). Granting schemes can be comprehended as a specific aspect of organization praxis at the juncture of state and professional expert logics. We advance that in this artistic production environment, in the absence of a market's logic, granting schemes and the interaction between professional evaluators and the state mediate the way in which a field is shaped.

Our case of the assessment of the performance of organizations in arts is exemplary for systems in which resource allocators such as the state and professional stakeholders do not espouse similar institutional norms nor share the same interests, similarly to what has been described elsewhere (Scott, 1987; Townley, 1997). Although public authorities set the directions for granting decisions, our analyses clearly show that the mere evaluations deviate from overt conditions. The quality of the artistic production is the major determinant for public support, which gives the impression that the experts behave as peers. Our case is also illustrative for an instance where the professions "lay successful claim to normative dimensions of political processes" (Townley, 1997: 280). This may not be a deliberate tactic by which proficiency overrules politics, but as a tacit practice its implications for

arts producers are not trivial. Diane Crane (1976) has advanced that those who control the key functions of the system set the cognitive and technical norms for innovation, on the basis of which they allocate rewards. Central to these resource allocation systems ("reward systems") are norms for the production of new products. In theory, in the absence of market selection, professionals dominating the resource allocation process have a major say about entry and exit, thus about the ecology of a field.

Performance appraisals such as the one exposed, have been described as powerful institutional myths, part of the rationale of what organizations need to be legitimate (Townley, 1997: 266). Consequently organizational performance could be largely contingent on these institutionalized norms and may be moderated by organizational legitimacy. If performance may be defined as the extent to which predetermined goals are achieved (Chow et al., 1994), a resource allocation system which is based upon performance standards that are not convenient nor transparent, or in which the evaluation is subject to underlying possibly conflicting value systems, reasonably matters to art producers that try to develop a sustainable business, balancing between mission and survival.

REFERENCES

Becker, H. S. (1974). Art as Collective Action. *American Sociological Review, 39*(6), 767-776.
Bielby, D. D., Moloney, M., & Ngo, B. Q. (2005). Aesthetics of Television Criticism: Mapping Critics' Reviews in an Era of Industry Transformation. *Research in the Sociology of Organizations, 23*, 1-43.
Bourdieu, P. (1983). The Field of Cultural Production, or the Economic World Reversed. *Poetics, 12*(4-5), 311-356.
Caves, R. E. (2000). *Creative Industries: Contracts between art and commerce*: Harvard University Press.
Chow, G., Heaver, T. D., & Henriksson, L. E. (1994). Logistics Performance: Definition and Measurement. *International Journal of Physical Distribution & Logistics Management, Vol. 24* (No. 1), pp. 17-28.
Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and Conducting Mixed Methods Research* (2nd ed.). Thousand Oaks, California: SAGE Publications, Inc.
Dimaggio, P. J., & Powell, W. W. (1983). The Iron Cage Revisited - Institutional Isomorphism and Collective Rationality in Organizational Fields. *American Sociological Review, 48*(2), 147-160.
Feder, T., & Katz-Gerro, T. (2012). Who benefits from public funding of the performing arts? Comparing the art provision and the hegemony-distinction approaches. *Poetics, 40*(4), 359-381.
Friedland, R., & Alford, R. R. (1991). Bringing society back in: Symbols, practices and institutional contradictions. . In W. W. Powell & P. J. DiMaggio (Eds.), *The new institutionalim in organizational analysis* (pp. 232-263). Chicago, Il.: University of Chicago Press.
Goodrick, E., & Reay, T. (2011). Constellations of Institutional Logics: Changes in the Professional Work of Pharmacists. *Work and Occupations, 38*(3), 372-416.
Gorman, G. E. (2007). The Oppenheim effect in scholarly journal publishing. *Online Information Review, 31*(4), 417-419.
Greenwood, R., Diaz, A. M., Li, S. X., & Lorente, J. C. (2010). The Multiplicity of Institutional Logics and the Heterogeneity of Organizational Responses. *Organization Science, 21*(2), 521-539.
Greenwood, R., & Hinings, C. R. (1996). Understanding radical organizational change: Bringing together the old and the new institutionalism. *Academy of Management Review, 21*(4), 1022-1054.
Guetzkow, J., Lamont, M., & Mallard, G. (2004). What is originality in the humanities and the social sciences? *American Sociological Review, 69*(2), 190-212.
Hadida, A. (2013 (forthcoming)). Performance in the Creative Industries. In C. Jones, M. Lorenzen & J. Sapsed (Eds.), *Oxford handbook of the Creative Industries* (pp. tba). Oxford: Oxford University Press.
Heinze, G., & Schemper, M. (2002). A solution to the problem of separation in logistic regression. *Statistics in Medicine, 21*(16), 2409-2419.
Hirsch, P. M. (2000). Cultural industries revisited. *Organization Science, 11*(3), 356-361.
Kassarjian, H. H. (1977). Content-Analysis in Consumer Research. *Journal of Consumer Research, 4*(1), 8-18.
Kolbe, R. H., & Burnett, M. S. (1991). Content-Analysis Research - an Examination of Applications with Directives for Improving Research Reliability and Objectivity. *Journal of Consumer Research, 18*(2), 243-250.
Krippendorff, K. (1980). *Content Analysis. An Introduction to Its Methodology.* California: SAGE Publications, Inc.
Levin-Rozalis, M. (2000). Abduction: A Logical Criterion For programme and Project Evaluation. *Evaluation, 6*(4), 415-432.
Marsh, H. W., Jayasinghe, U. W., & Bond, N. W. (2008). Improving the peer-review process for grant applications - Reliability, validity, bias, and generalizability. *American Psychologist, 63*(3), 160-168.

Meyer, J., & Rowan, B. (1991). Institutionalized organizations: formal structure as myth and ceremony. In W. Powell & P. DiMaggio (Eds.), *The new institutionalism in irganizational analysis* (pp. 41-62). Chicago: University of Chicago Press.
Oakes, L. S., Townley, B., & Cooper, D. J. (1998). Business planning as pedagogy: Language and control in a changing institutional field. *Administrative Science Quarterly, 43*(2), 257-292.
Peterson, R. A., & Anand, N. (2004). The production of culture perspective. *Annual Review of Sociology, 30*, 311-334.
Reay, T., & Hinings, C. R. (2005). The recomposition of an organizational field: Health care in Alberta. *Organization Studies, 26*(3), 351-384.
Reay, T., & Hinings, C. R. (2009). Managing the Rivalry of Competing Institutional Logics. *Organization Studies, 30*(6), 629-652.
Scott, W. R. (1987). The Adolescence of Institutional Theory. *Administrative Science Quarterly, 32*(4), 493-511.
Shatz, D. (2004). *Peer review: a critical inquiry.* Lanham: Rowman & Littlefield.
Strijbos, J. W., Martens, R. L., Prins, F. J., & Jochems, W. M. G. (2006). Content analysis: What are they talking about? *Computers & Education, 46*(1), 29-48.
Suchman, M. C. (1995). Managing Legitimacy - Strategic and Institutional Approaches. *Academy of Management Review, 20*(3), 571-610.
Townley, B. (1997). The institutional logic of performance appraisal. *Organization Studies, 18*(2), 261-285.
Turbide, J., & Laurin, C. (2009). Performance Measurement in the Arts Sector: the Case of the Performaing Arts. *International Journal of Arts Management, 11*(2), 56-70.
Venkatraman, N., & Ramanujam, V. (1986). Measurement of Business Performance in Strategy Research - a Comparison of Approaches. *Academy of Management Review, 11*(4), 801-814.
Voss, G. B., Cable, D. M., & Voss, Z. G. (2000). Linking organizational values to relationships with external constituents: A study of nonprofit professional theatres. *Organization Science, 11*(3), 330-347.
White, H. C., & White, C. A. (1993 (1965)). *Canvases and Careers: Institutional Change in the French Painting World*. Chicago: University of Chicago Press.