Research Evaluation, 2022, 1–13
https://doi.org/10.1093/reseval/rvac017
Article

OXFORD

Evaluation of the arts in performance-based research funding systems: An international perspective

Kamila Lewandowska [1,*], Emanuel Kulczycki [2] and Michael Ochsner [3]

[1]The Aleksander Zelwerowicz National Academy of Dramatic Art in Warsaw, ul. Miodowa 22/24, 00-246 Warsaw, Poland, [2]Scholarly Communication Research Group, Adam Mickiewicz University in Poznań, ul. Międzychodzka 5, 60-371 Poznań, Poland and [3]FORS, Swiss Center of Expertise in the Social Sciences, University of Lausanne, Géopolis, 1015 Lausanne, Switzerland

*Corresponding author. Email: kamila.lewandowska@e-at.edu.pl

Abstract

This article provides a comprehensive analysis of the evaluation of the arts within performance-based research funding systems (PRFSs). Previous literature on PRFSs has overlooked the arts and focussed primarily on outputs in relation to the sciences and humanities. We develop a typology of how artistic outputs are evaluated within 10 countries' PRFSs, operating in Australia, the Czech Republic, Italy, Lithuania, New Zealand, Poland, Portugal, Slovakia, Spain, and the UK, and identify three different types of artistic evaluation systems. The study compares evaluation methods and provides a classification of quality criteria used by evaluation panels. We conclude with a discussion of the challenges specific to different types of systems.

Key words: performance-based research funding system; research evaluation; the arts; artistic research, research quality criteria

1. Introduction

Performance-based research funding systems (PRFSs) have been studied extensively (Hicks 2012), but we still know surprisingly little about the evaluation of the arts in the context of PRFSs. Although evaluation in artistic disciplines has been examined from different perspectives—for example, scholars have analysed assessments of practice-based doctorates (Dally et al. 2004; Kroll and Webb 2012; Nilsson, Dunin-Woyseth and Janssens 2017), grant applications (Hellström 2010) as well as formal and informal university-level assessments which impact scholars' academic lives (e.g. Wilson 2016), those perspectives do not include PRFS evaluations.

In this article, we intend to address this gap by shedding light on the evaluation of artistic outputs within different PRFSs. Specifically, we follow three sequential research questions. First, we review systematically the literature and identify the knowledge gaps linked to the evaluation of the arts within PRFSs. Second, we analyse 10 national PRFSs and develop a typology of how the arts are evaluated using the information the countries provide on performance-based research evaluation systems. Finally, we analyse how those systems evaluate the arts, focussing particularly on 'artistic outputs' as PRFSs are linked to outputs submitted for the evaluation exercise. We also compare the methods used to evaluate artistic outputs and develop a typology of evaluation criteria for the arts in use in those systems.

Thus, the overall aim of this study is to improve understanding of PRFSs by focussing on the unexplored area of artistic disciplines. This is relevant for several reasons. First, learning how the arts are evaluated within research funding systems helps better understand some more general contemporary changes in the higher education sector. One of those current trends is academization of vocational education—the emphasis on making vocational education more 'scientific' through linking professional training to research (Ek et al. 2013). Academization is closely connected to large-scale higher education reforms such as the inclusion of post-secondary programmes into the system of higher education and absorption of specialized vocational schools into university structures (e.g. Rust, Mottram and Till 2007). As a result of those reforms, there have been increasing demands that academics in the fields hitherto dedicated to

professional education—such as the arts (Johansson and Georgii-Hemming 2021), nursing (Laiho 2010), or social work (Kessl et al. 2020)—would be more involved in research activities and adapt to university research cultures and standards. This study contributes to knowledge on academization by showing how such demands are operationalized within, and executed through, national research funding systems.

Second, an analysis of the evaluation of the arts within systems designed to assess research outputs sheds light on different but overlapping concepts of value and is therefore helpful from the perspective of the sociology of worth. The introduction of the arts into the area of research evaluation has brought into co-existence two different modes of evaluation: the artistic mode and the research mode. Those two modes are linked with potentially competing or incompatible values; for example, while a systematic, methodological approach is highly valued in research disciplines (Hug, Ochsner and Daniel 2013), it is rarely used as a standard in the arts where creativity and inspiration play much greater role than methodological transparency (Elliott 2011; Croft 2015). Therefore, in PRFSs evaluating artistic outputs, different 'orders of worth' (Boltanski and Thévenot 2006)—distinctive repertoires of evaluation, each with its own set of evaluative criteria and related conception of the 'common good'—operate within a single domain. Studying artistic evaluation within PRFSs can thus improve our understanding of how evaluation systems accommodate and integrate different evaluative principles.

This article is structured as follows. First, we define the scope of the study by describing the inclusion criteria for the PRFS and present a review of research on the evaluation of the arts in PRFSs. We then systematize artistic evaluation in PRFSs and suggest a typology of how artistic outputs are integrated into evaluations within PRFSs. Finally, we analyse how the arts are evaluated in PRFSs with a special focus on methods and evaluation criteria.

2. The scope of the study

Evaluations and classifications of disciplines vary widely across countries and contexts. We restrict our study to the field of art (the arts), which we define as disciplines including the practice of art, sometimes called the 'performing and visual arts', such as dance, theatre, music, creative writing, or design, but exclude predominantly analytical disciplines such as art history, musicology, or literature studies. By 'artistic outputs', we mean outputs of artistic (research[1]) practices, which take the form of artefacts, compositions, live performances, exhibitions etc. and which are considered as eligible outcomes within PRFSs. We restrict our analysis to national performance-based evaluation systems. Particularly, we define PRFSs based on criteria of eligibility proposed by Hicks (2012) and adapted for our purposes:

- Evaluations must be part of a **PRFS**. Evaluations of artistic productions taking place outside the context of those systems are excluded. Consistent with Hicks's definition of PRFSs, a PRFS must be a national system, evaluations must be *ex-post* and governmental allocation of funding must depend on the evaluation outcome. Only evaluations of artistic outputs of academic staff members are included (e.g. student achievements are excluded).
- **Artistic outputs** must be evaluated. Evaluations of traditional research outputs (publications) or non-traditional research outputs that are not results of artistic activity (e.g. a research report, software) are excluded.
- The **quality** of outputs is evaluated. Evaluations of research impact case studies, research income, research management, or other aspects of institutional research activity are not analysed in this article.

To select the country systems relevant for the purposes of the study, the authors first derived a list of research evaluation systems from the following pertinent publications: Hicks (2012), Jonkers and Zacharewicz (2017), and Zacharewicz et al. (2019). Second, documentation on evaluation exercises performed in 2018 or later was retrieved from the websites of evaluating agencies and reviewed. The authors identified PRFSs which included evaluation of artistic outputs and conducted a review of the main characteristics of PRFS evaluations. This resulted in a selection of 10 PRFSs operated in Australia,[2] the Czech Republic, Italy, Lithuania, New Zealand, Poland, Portugal, Slovakia, Spain, and the UK.

3. Literature on evaluation of the arts within PRFSs

To identify literature on, and gaps of knowledge about the evaluation of the arts within PRFSs, we conducted a systematic search of seven academic research databases: Dimensions, ERIC, Google Scholar, Microsoft Academic, Science Direct, Scopus, and Web of Science. Each bibliographic database had a different search mode and our search parameters had to be modified accordingly. We used relevant keywords and Boolean operators (e.g. 'artistic research' OR 'arts-based research' OR 'artistic disciplines' OR 'performing arts' OR 'visual arts') with the additional search terms ('research evaluation' OR 'research assessment'). The three inclusion criteria were as follows:

1. Research publications in English including articles, books, and book chapters available in at least one of the seven mentioned databases.
2. Publications considering the evaluation of the arts within national performance-based research evaluation systems.
3. Publications published between 2001 and 2020. The period of 20 years was chosen because it catches a wide range of publications and allows to present the evolution of approaches to the evaluation of the arts.

Database searches were augmented with hand searches of journals relevant to the topic, such as *Research Evaluation*, *Assessment & Evaluation in Higher Education*, and *International Journal of Education and the Arts*. Reference lists of the already included studies were screened to identify possibly relevant papers. We contacted The European League of Institutes of the Arts and asked for literature recommendations.

The searching process is summarized in Figure 1. The process led to 1,531 possibly eligible studies. After removing duplicates and screening titles and abstracts for the three inclusion criteria mentioned above, 92 studies remained and were analysed one by one. The vast majority of those 92 studies dealt with the evaluation of artistic research but only 17 studies concerned PRFS evaluations. We list the 17 studies focusing on PRFSs in the reference list indicated by a star (*).

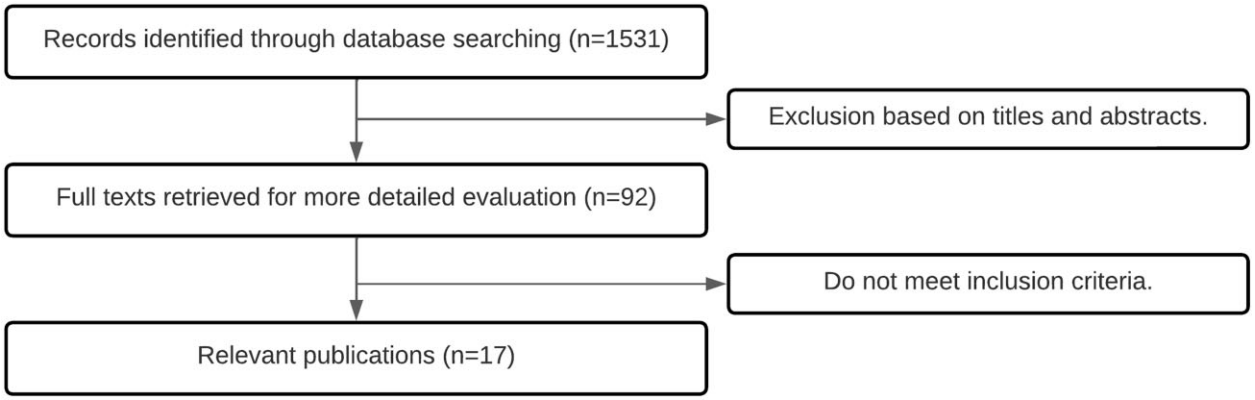

Figure 1. Flow diagram of the literature review.

This result draws attention to the fact that the evaluation of artistic outputs within PRFSs is an overlooked and under-researched topic. In particular, we identified three relevant gaps of knowledge. First, the majority (14 out of 17) of studies focus on the Anglo-Saxon country context (Australia and the UK, i.e. Green 2006; Rust, Mottram and Till 2007; Bennett, Wright and Blom 2009; Wright, Bennett and Blom 2010; Wilson 2011; Bennett and Franzmann 2013; Harrison and Draper 2014; Barwick and Toltz 2017; Butt 2017; Schippers, Tomlinson and Draper 2017; Brook 2018; Wilson 2018; McKee 2020; Toltz 2020). Second, while the studies discuss many important aspects of PRFSs—for example, the process of integrating the arts into PRFSs (e.g. Green 2006; Rust, Mottram and Till 2007; Bennett and Franzmann 2013; Butt 2017; Schippers, Tomlinson and Draper 2017; Sivertsen 2018), the ways universities organize data collection and submission (Harrison and Draper 2014; Barwick and Toltz 2017) and the challenges evaluation systems create for artist-academics (e.g. Bennett, Wright and Blom 2009; Wright, Bennett, and Blom 2010; Wilson 2018)—they generally do not provide a detailed analysis of evaluation procedures and criteria. Third, studies comparing different countries' models are entirely lacking.

4. PRFSs taking into account artistic outputs: an overview

Using the results of the literature review and by analysing the documentation of PRFSs, we identified 10 countries that include evaluation of artistic outputs in their PRFSs. The identified systems are presented in Table 1. Countries, the name of the PRFS (or agencies, if the name of the system is not specified), types of artistic outputs (how artistic outputs are called or defined in the evaluation documents), the arts (research), fields, and sub-fields (how the arts field is called, categorized and subdivided within the classification of research and development fields used in a specific PRFS) are listed in Table 1.

The categorization of artistic outputs and the accommodation of the arts within the remit of research funding differ strongly across countries. We distinguish two dimensions that are relevant for the analysis of the evaluation of arts in PRFSs. First, a PRFS might evaluate artistic research or artistic activity. Second, types of outputs are treated equally, or scores depend on the type of outputs.

4.1 Dimension 1: artistic research versus artistic activity

In some country systems, artistic outputs submitted for evaluation are assessed as 'research outputs'. Only artistic outputs that qualify as research are considered eligible for evaluation. Those systems are based upon a distinction between artistic research and professional art practice—only the former can be assessed within the PRFS and contribute to the final evaluation score. This practice has its foundation in theories that view some artistic productions or practices as a form of research because the artistic productions contribute to knowledge through a process of investigation; at the same time, not all art has a research component and should be included in research funding systems (Strand 1998; Biggs and Karlsson 2010). We call these 'artistic research' systems (see Figure 2). Other PRFS designs are based upon a distinction between research and artistic activity. Artistic disciplines are included in the assessment process, but artistic outputs are not required to meet the criteria of research and are evaluated on the basis of their artistic merit. We call these 'artistic activity' systems.

This distinction is relevant for the evaluation of the visual and performing arts within PRFSs because they have different practical implications for the evaluation exercise, as we will describe in more detail below.

4.2 Dimension 2: treatment of different types of outputs

Whereas all PRFSs define different types of outputs eligible to the evaluation exercise, some PRFSs explicitly treat all types of outputs equally. They start from the premise that the quality of research does not depend on the type of output or that the type of output is not a consequence of the quality of research. Rather, all types of outputs can be of higher or lower quality or reflect better or less good research. We call these 'equality of outputs types' systems (see Figure 2). Other PRFSs attribute different scores to different types of outputs. The underlying idea is that some outputs take more time and effort to produce than others. For example, preparing a solo exhibition takes more time and more effort than participating in a group exhibition. It also implicitly assumes that the effort is somehow linked to the audience it can reach. Furthermore, a differential scoring of outputs leads to transparency as everything is more clearly defined. It may be more practical in cultural contexts characterized by lower trust in public experts, where such transparency of evaluation criteria and methods is particularly valued. We call these systems 'differential score of output types' systems.

Table 1. Performance-based research funding systems in which artistic outputs are evaluated

Country	System/Agency	Output category of artistic outputs	Field and sub-fields in which performing arts are classified
Australia	Excellence in Research for Australia (ERA)	'Non-traditional research outcomes (NTRO)' (Australian Research Council (ARC) 2017)	Studies in creative arts and writing: • Art theory and criticism • Film, television, and digital media • Journalism and professional writing • Performing arts and creative writing • Visual arts and crafts • Other studies in creative arts and writing
Czech Republic	Evaluation of research, development and innovations ('R&D&I') (2017+)	'Other [research] results: artistic research' (Government of the Czech Republic 2018)	Humanities and the arts: • Arts (arts, history of arts, performing arts, and music)
Italy	Evaluation of Research Quality (VQR)	'Other scientific outputs: Compositions; Drawings; Design works; Performances (art, theatre, music)...' (Italian National Agency for the Evaluation of Universities and Research Institutes (ANVUR) 2020)	Ancient History, Philology, Literature, and Art History (i.e. performing arts are evaluated with their 'closest relative' in the humanities)
Lithuania	Research and Experimental Development of Lithuanian Research and Study Institutions and Evaluation of Artistic Activities	'Works of art' (Ministry of Education and Science of the Republic of Lithuania (SMM) 2017)	Art, design, architecture, media art, theatre, cinema, dance, music or interdisciplinary art
New Zealand	Performance-Based Research Fund	'Non-traditional research output types' (New Zealand, Tertiary Education Commission 2018)	Creative and performing arts: • Design • Music, literary arts, and other arts • Theatre and dance, film and television, and multimedia • Visual arts and crafts
Poland	Evaluation of quality of scientific and artistic activity	'Artistic achievements' (Poland, Ministry of Science and Higher Education 2019, 2020)	The Arts: • Theatre and film • Music • Fine arts and conservation
Portugal	The Portuguese Science and Technology Foundation (FCT)	No specific category for artistic outputs (Portugal, Portugal, Foundation for Science and Technology (FCT) 2018)	Arts and humanities:. • Arts and Design, Artistic and Musical Development
Slovakia	The Slovak Accreditation Agency for Higher Education	'Artistic activities' (Slovak Accreditation Agency for Higher Education 2018)	Artistic group: Arts
Spain	Sexenio—Retrospective research evaluation system (RES)	No specific category for artistic outputs (all outputs must meet the criteria of research) (Spain, National Commission for Evaluation of Research Activity (CNEAI) 2020)	History, Geography, and the Arts: • Fine arts • Music
UK	Research Excellence Framework (REF)	No specific category for artistic outputs (all outputs must meet the criteria of research) (Research Excellence Framework (REF) 2019b)	• Art and Design: History, Practice, and Theory • Music, Drama, Dance, Performing Arts, Film, and Screen Studies

4.3 Typology of PRFSs taking into account artistic outputs

Combining the two dimensions leads to a four-field scheme with which we can classify PRFSs according to how they include the arts into their system (Figure 2). On the *x*-axis, we differentiate systems that evaluate artistic research versus systems evaluating artistic activity. On the *y*-axis, we distinguish systems applying the principle of 'equality of output types' from systems based on the principle of 'differential score of output types'.

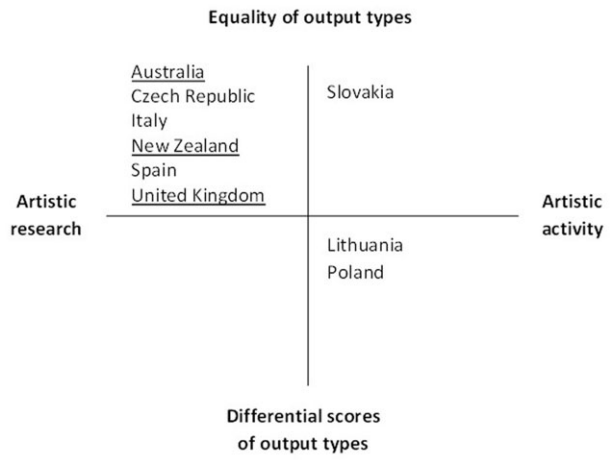

Figure 2. Two dimensions of artistic outputs' evaluations within PRFSs. Note: Countries are sorted in alphabetical order in each cluster. Underlined names represent countries in which research statements are used.

Theoretically, four types of PRFSs' way of including the arts are possible, but we only identified three types in our selection of PRFSs. First, we have PRFSs evaluating 'artistic research' and applying the principle of 'equality of output types'. Australia, the Czech Republic,[3] Italy, New Zealand, Spain, and the UK belong to this type. Second, we identified one PRFS evaluating 'artistic activity' and applying the principle of equality of output types', Slovakia. Third, two PRFSs evaluate 'artistic activity' and apply the principle of 'differential score of output types', Lithuania and Poland.

Having systematized how the arts are included in PRFSs, we can use the types to describe how the arts are evaluated in PRFSs. First, we will focus on artistic outputs and their classification in PRFS, then we will describe the evaluation methods used to evaluate artistic outputs and then we will analyse the evaluation criteria used to evaluate artistic outputs.

5. Evaluation of artistic outputs within PRFSs

5.1 Types of outputs

Contemporary art encompasses a substantial diversity in format, technique and medium. This is why in all selected PRFSs, a wide diversity of outputs can be submitted for evaluation. Eligible outputs may include original works of art (e.g. artefacts, objects, compositions, designs), live performances, exhibitions, portfolios, etc. However, the way how artistic outputs are classified depends on how PRFSs evaluate them. In systems applying the principle of 'equality of output types', it is not necessary to have a detailed list of types of outputs. Accordingly, art units in some countries are allowed to submit also the achievements not fitting directly into the pre-defined outcome categories listed in the guidelines. The REF system, for example, even recognizes explicitly that one output may encompass different output types, for example, an artefact can also be a subject of an exhibition. Thus, the REF accepts submissions that 'can include but are not limited' to the pre-defined output forms.

The PRFSs based on the principle of 'differential score of output types' on the other hand provide a closed catalogue of outputs eligible for evaluation because each type of output needs to be assigned a score or range of scores. Poland has probably the most stringent model based on a detailed classification of artistic forms. This approach may seem less appropriate for the evaluation of contemporary art, which transgresses borders of artistic formats and is difficult to capture in traditional artistic divisions (Lundén and Sundén 2015). At the same time, a pre-determined catalogue of art forms and activities seems more practical for systems that, like the Polish PRFS, combine peer review-based assessment with the evaluation based on quantitative metrics. In Poland, faculties obtain points for submitted artistic outputs and the number of points depends on the output type (see Supplementary Table S1 presenting eligible artistic outputs in the Polish PRFS).

To show the differences in how to include different artistic output types into PRFSs, we present in more detail how the British REF—an 'artistic research' system based on the principle of 'equality of output types'—and the Polish PRFS—an 'artistic activity' system based on the principle of 'differential scores of output types'—define artistic outputs and integrate them into their system. Accordingly, Table 2 illustrates differences in defining and scoring outputs between the British REF and Polish PRFS (please note that the table presents just a few examples of output types; for complete lists of outputs see Supplementary Tables S1 and S2). In the British REF, the guidelines contain examples of outputs, but a submission does not have to represent one of the listed artistic formats even though the definitions are quite detailed. The same definitions apply for all art disciplines. In Poland, eligible outputs are precisely defined and each output is assigned a specific range of points. Submissions not fitting into one of the listed output types are not eligible. Moreover, in the Polish PRFS, output types are defined per artistic disciplines. Lists differ thus between 'music', 'film and theatre' and 'fine arts and art conservation' faculties (see Supplementary Table S1 for the output definitions for 'music' faculties).

In 'artistic research systems', besides evaluating the content of non-traditional research *outputs*, the research *process* that led to the output needs to be specified and verified that it actually includes aspects of research. To make the research component evident authors must attach a 'research statement' explaining the investigatory dimensions of the submission. The reason behind this requirement is that the research component of an artistic output may not be evident within the product itself. The product of artistic research is not necessarily different from the product of non-research art practice—it is the research process that makes the former different from the latter (Strand 1998; Vanlee and Ysebaert 2019) and therefore needs to be explained to the evaluators.

For example, submissions to the UK REF are expected to be accompanied by a 300-word description of the research process (research questions, methodologies, and referenced literature), research insights (findings and outcomes) and dissemination (where the insights were effectively shared, see Research Excellence Framework (REF) 2019b). The Australian ERA and the New Zealand PBRF require descriptions addressing research background, contribution, and significance of outputs (Australian Research Council (ARC) 2017; New Zealand, Tertiary Education Commission 2018). In other systems, research statements may not be required but non-traditional outputs must meet the definition or criteria of research specified in the PRFS guidelines.

Slovakia, as the only system evaluating 'artistic activity' with 'equality of output types', follows indeed the characteristics of the two dimensions mentioned above. Types of output do not play a role and, as artistic activities are evaluated, no research statement has to be submitted. However, detailed information is asked to guarantee a holistic evaluation. This not only includes information

Table 2. Examples of output types and scoring in the Polish PRFS and the British REF

Polish PRFS		British REF	
Output type (examples)	Scoring	Output type (examples)	Scoring
An original music composition for a large ensemble (more than 15 performers) made available in the public domain through performance, recording, broadcasting, or published score A leading part in a musical, ballet, or dance performance	200 points (excellent) 100 (significant for the discipline) 50 (other)	Composition: an original published/publicly available score, first performance, or first recording by a record label of a musical composition. This can include (but is not limited to): compositions created while being played, e.g. electronic compositions, jazz improvization; published/publicly available score; recordings; sound component of a film or video, lyrics, multimedia composition; commissioned works; combinations or developments of the above	Scores not attributed to output types. Each output listed in a submission is assessed against the quality levels: 4*, 3*, 2*, 1*, or 'unclassified'
An original music composition; a recording, broadcasting, or published score of a music, electronic, or multimedia composition for a smaller ensemble (by soloists or chamber ensembles up to 15 performers) made available in the public domain through performance, recording, broadcasting or published score; A supporting part in a musical, ballet, or dance performance;	200 points (excellent) 75 (significant for the discipline) 40 (other)		
Participation as an art director or manager in a music competition, festival, or series of concerts;	200 points (excellent) 50 (significant for the discipline) 25 (other)		

on authorship, publication date and location etc. but also citations or a 'brief annotation with the contextual information on the impact of the output on socio-economic practice' that should be factual and verifiable (Slovak Accreditation Agency for Higher Education 2018: 40).

5.2 Evaluation methods

Peer review is the most widely used method of artistic outputs' assessment. Even though evaluation of artistic productions is typically a qualitative and descriptive process, when used in decision-making, the qualitative assessments often need to be translated into quantitative measures (Lewandowska and Smolarska 2020). All analysed PRFSs utilize discipline-based expert panels to do so, although disciplinary arts fields vary, depending on local classifications of research fields. Table 1 presents the disciplinary scope of panels assessing artistic outcomes. In some PRFS, artistic activity is evaluated within the humanities field (the Czech Republic, Italy, and Spain), whereas in others art is a separate field (Australia, Lithuania, New Zealand, Poland, Slovakia, and the UK). A stronger autonomy of art as a separate discipline falls usually together with more detailed and tighter evaluation guidelines, including the specification of output types and indicators that are (or are not) applicable in this field. Note that, by definition, in all 'artistic activity' systems, and especially in those using a differential scoring per output type, art is a separate discipline and enjoys some autonomy regarding output types (i.e. has its 'own' output types).

Discipline-based assessments carry a risk of underestimating research that crosses disciplinary boundaries (Laudel 2006). Some PRFSs recognize this threat and introduce solutions facilitating interdisciplinary peer review. In the REF, for example, submissions are cross-referred to other panels for advice, if the main panel does not consider it contains the relevant expertise to assess the quality of an output. The evaluated unit may request that specific parts of submissions should be cross-referred to another disciplinary panel. The Australian ERA allows assignment of a research output to four different disciplinary sub-fields which reflect their content and indicate the disciplinary expertise needed to assess their quality. In Italy, panels employ two external independent assessors if the disciplinary expertise necessary for evaluating a specific output is lacking.

Besides commonalities, the PRFSs differ in how artistic outputs are evaluated. In 'artistic research' systems using 'equal treatment of output types', the descriptive evaluation of each output is translated into a quantitative measure using five or seven quality levels; typically, higher level indicates higher quality except for the Czech system where five is the highest score and one is the lowest. In the UK, panels create a unit's overall quality profile by calculating the percentage of outputs listed in a submission that are assigned at each quality level. In New Zealand, outputs are scored using a zero- to seven-point scale and the assessment is a two-stage process: first, the preparatory scores are determined individually and then collectively by two selected panel members and, second, calibrated panels' scores are determined by the whole panel and a quality category is assigned. The documentation of outputs submitted must provide sufficient information that allows panel members to personally judge the artistic or research value of outputs. It is expected that panels will not rely solely on the 'proxies for quality', such as venue prestige or peer recognition, and carefully examine a proportion of outputs so that the assessment of quality is based on the quality of the output itself. The REF guidelines explicitly instruct panels to 'ignore any additional material that includes evaluative commentary on the perceived quality of a research output' (Research Excellence Framework (REF) 2019b: 60), such as critical art reviews.

In 'artistic activity' systems using 'differential scores for output types' in our case Poland and Lithuania, scores depend on the formal classification of the outputs. However, the two countries differ

considerably in how this is done. In Lithuania, some outputs including works of art, artistic performances, awards and artistic research are assigned with a range of Points (2–7), whereas others, such as art reviews or curation of art shows are awarded only one point. In Poland, each output can be assessed as 'excellent', 'significant for the discipline', or 'other' and receive a number of points corresponding to each of those quality levels. However, the numbers of points vary depending on output type, for example, a leading part in a ballet performance can receive 200 (excellent), 100 (significant), or 50 points (other), whereas a supporting role—200, 75, or 40 points (see Table 2). Reviewing panels determine the quality level of each output—however, due to the pre-determined scoring, reviewers have less freedom in assigning points than panels in 'equal treatment of output types' systems. Panel evaluations are also more formal and less qualitative, as panellists do not have access to the actual content of outputs—they are not provided with audio, video, etc.—and evaluate them primarily for the received awards and recognition of the artistic community (see the definition of 'excellent' work at the bottom of Supplementary Table S1). Again, Slovakia as the sole representative of an 'artistic activity' using 'equal treatment of output types' system combines the characteristics of the two systems described above. Similar to the 'artistic research' using 'equal treatment of output types' systems, outputs are graded according to a five-level qualitative statement resulting from a holistic evaluation. From the single scores, each unit is given a quality profile by calculating the percentage of outputs listed in a submission that are assigned at each quality level. The final overall score is calculated by multiplying the percentage with the score (Slovak Accreditation Agency for Higher Education 2018: 45).

5.3 Evaluation criteria

As PRFSs tend to emphasize transparency of evaluation methods (Hicks 2012) and comparability of assessment results, official criteria are established and publicly announced before the assessment exercise and panels are required to follow them strictly. In the following, we systematize criteria for the evaluation of artistic outputs in PRFSs. To derive the criteria, we examined official documents (evaluation regulations, submission guidelines, etc.) where the criteria are defined. We found that there are two distinct types of criteria: formal criteria and quality criteria. Formal criteria are the requirements that a submission must satisfy to be eligible for the assessment process. They are explicitly listed in the evaluation regulation and submission guidelines. Quality criteria are used to guide the quality assessment process and assignment of scores to artistic outputs carried out by expert panels. They are either explicitly listed (e.g. in statements such as: 'the panels will assess the quality of research outputs in terms of their "originality, significance, and rigour"') or provided more implicitly by indicating what qualities characterise outputs at each 'quality level', for example:

> For an A to be assigned it would normally be expected that the EP contains evidence of research output of a world-class standard and research-related activity that shows a high level of peer recognition and esteem within the relevant research subject area and indicates a significant contribution to the New Zealand and/or international research environments (...). (New Zealand, Tertiary Education Commission 2018: 32)

To systematize quality criteria, the fragments of texts describing outputs' qualities were derived from the documents and coded manually. An inductive, *in vivo* data coding was used and two rounds of coding were conducted. The first round was focussed on summarizing bits of text using words (code names) taken from the text itself, such as 'peer recognition', 'significant contribution to research'. The aim of the second round was to re-examine and reduce the number of codes by ordering them into more general categories (quality criteria).

5.3.1 Formal criteria

The formal criteria are clearly defined in the evaluation regulation and submission guidelines to facilitate the submission process. Typically, outputs must be (co-)produced or (co-)authored by one or more members of the evaluated institution's research staff and disseminated in the public domain during the assessment period.

Formal criteria also include a minimum and a maximum number of outputs that can be submitted for evaluation. For example, in the UK, the total number of outputs must equal 2.5 times the summed full-time equivalent of the unit's staff; rounding to the nearest integer is applied (see Research Excellence Framework (REF) 2019a: 8). In addition, each staff member with a contract of employment of 0.2 FTE (full-time equivalent) or greater must be a (co-)producer or (co-)author of at least one output but a maximum of five outputs may be attributed to an individual staff member. In New Zealand, the submission may contain up to 16 research outputs, including a maximum of 4 outputs nominated for peer assessment (New Zealand, Tertiary Education Commission 2018: 13). In Slovakia, institutions may submit 5 outputs for each eligible researcher and up to 25 per unit (Slovak Accreditation Agency for Higher Education 2018: 39). In Lithuania, there is a limit of 30% of the outputs that can be submitted to the evaluation (Ministry of Education and Science of the Republic of Lithuania (SMM) 2017).

5.3.2 Quality criteria

The quality criteria are developed through consultation with experts representing universities and other research institutions. The consultation procedures vary depending on the PRFS model. For example, in the REF, four main expert panels representing different research fields work together to develop a combined set of quality criteria. These can be supplemented by discipline-specific, more detailed criteria which nevertheless must be read alongside the generic criteria. In 'artistic research' systems, the same generic criteria apply to all research disciplines and fields, including the field of art. In 'artistic activity' systems, specific criteria are developed for artistic disciplines. In Poland, the quality criteria for artistic evaluations differ from those used in other disciplines, where the scoring of outputs is based primarily on the Ministry's ratings of journals and publishers. Instead of a combined, cross-disciplinary set of quality standards, the Polish PRFS incorporates criteria developed by the representatives of higher art education institutions.

The analysis of documents allowed identifying 12 higher-order quality criteria. Table 3 shows which criteria are more universally applied across different systems and which are used only in some PRFSs. *Significance* and *originality* are quality criteria present in all PRFSs. However, depending on the type of PRFS the scope of *significance* and *originality* changes. In 'artistic research' systems, significance is limited to research, whereas in 'artistic activity' systems, significance for art is evaluated. In a similar vein, only the 'artistic research' systems evaluate outputs in terms of their *contribution to knowledge/understanding* in the research field. They emphasize a broad understanding of research that goes beyond theoretical

Table 3. Quality criteria used in peer evaluations of the arts within PRFSs

	Czech Republic	Italy	Portugal	Spain	UK	Australia	New Zealand	Slovakia	Lithuania	Poland
Contribution to knowledge/understanding	✓	✓	✓	✓	✓	✓	✓			
Significance for research	✓	✓	✓	✓	✓	✓	✓			
Significance for art								✓	✓	✓
Originality (extrinsic)	✓	✓	✓	✓	✓	✓	✓	✓	✓	
Originality (intrinsic)						✓	✓		✓	
Rigour		✓			✓		✓	✓		
Creative/intellectual context					✓	✓	✓			
World-class level	✓		✓		✓		✓	✓	✓	
International exposure								✓	✓	✓
Peer recognition				✓	✓	✓	✓	✓	✓	✓
Scale of work										✓
Output type									✓	✓

development or empirical findings and includes creative (artistic) investigations. Contribution to 'understanding' is thereby on a par with the contribution to 'knowledge', to take into account the forms of cognition more typical for art (Riley 2019). *Originality*, one of the most used criteria in research evaluations (Guetzkow, Lamont and Mallard 2004), also comes in two forms. Their link to the PRFS system is, however, a little weaker. Originally understood as a new contribution to art or research (*extrinsic originality*) prevails in 'equal treatment of outputs' systems (but Lithuania as a 'differential treatment of outputs' system also uses it). Originality understood as a new contribution to an earlier version of the output (*intrinsic originality*) applies to the two Oceanic systems, Australia and New Zealand, as well as Lithuania. This criterion relates to the fact that a creative output that is repeatedly disseminated may evolve over time with a resulting change in quality and impact. Therefore, a unit can count repeated exhibitions or events as multiple outputs if they introduce a new component or interpretation to the work. In Lithuania, however, repeated disseminations cannot be awarded a full number of points but are multiplied by a so-called 'uniqueness coefficient' of 0.8, 0.5, or 0.4.

Another well-known criterion in research evaluation is *rigour*, which refers to intellectual coherence and integrity. It is surprising that it is found only among four PRFSs in our sample. It is only relevant in 'equal treatment of output types' systems as the classification of outputs in 'differential treatment of output types' includes the assumption that rigour is linked to output type. We expected that rigour would be present only in 'artistic research' system but found it also in Slovakia. As a similar criterion as *rigour*, we identified the criterion *creative or intellectual context* of the artwork to account for the specificities of artistic research in some 'artistic research' systems. It refers to the creative references upon which the work draws and the familiarity in the research statement with the current state of knowledge in the artistic discipline.

The *international dimension* of outputs is evaluated in most PRFSs. However, it also comes in two forms, *world-class level* and *international exposure*. The criterion *world-class level* is not dependent on the type of PRFSs. In general, PRFSs use this criterion to assess the quality standard, not a geographic scale or research scope of submissions. The term *world-class* means that the output qualifies as one of the best within its discipline and it is assumed that themes of primarily local scope can be of world-class standard. In contrast, the 'artistic activity' PRFSs (Lithuania, Slovakia, and Poland) apply the *international exposure* criterion—that is, take into account if the output was realized or presented abroad.

In most systems, *peer recognition* is considered in the assessment process. Outputs that have received validation from professionals in the field—for example, obtained awards in significant competitions, were publicly presented in prestigious institutions or at prominent festivals, or selected for inclusion in renowned exhibitions—are scored more highly than outputs that have not received considerable attention from field-based experts. In New Zealand, the 'quality-assurance' of an output—the fact that it has been subject to expert review before entering the public domain—plays an important role in PRFS evaluations. It includes not only the traditionally scientific forms of peer review but also the processes of validating artistic outputs undertaken by museums, theatres, broadcasters, etc. The utilization of the already existing, professionally embedded systems of artistic reviewing in PRFS assessments is generally supported by art professionals (Blythe 2018). However, the extent of this utilization varies: while in 'differential treatment of output types' systems prizes, venue prestige, or festival selections are treated as main indicators of output quality, in 'equal treatment of output types' systems, where panels are instructed to undertake a quality assessment of the actual work (see Section 5.2), they are expected to play a supplementary role.

There are some criteria that appear only in a few PRFSs. *Scale of work*, for example, a theatre performance of at least 70-min duration is rated more highly than a performance of <70 min, is used in only one country: Poland. However, in the Polish system the role of this indicator has diminished over the last years and since the enactment of new regulations in 2019, few output types are evaluated in terms of scale (see Supplementary Table S1). The use of this criterion in PRFSs is not universally accepted and there is a lack of consensus among experts whether the size or duration of an artwork is related to its artistic or academic quality (McKee 2020). However, given that the scale of work is usually more easily measurable than, say, its aesthetic significance, this criterion may be valued for its practicality and transparency.

Finally, in 'artistic activity with differential treatment' systems, the scoring of outputs depends on *output type*—for example, a music composition for a symphony orchestra is scored differently than a curation of an exhibition.

6. Discussion

Previous literature on PRFSs has overlooked the arts and focussed primarily on outputs in relation to the sciences and humanities. Our

article shows that there is a wide variety of approaches within national PRFSs regarding the assessment of artistic research and activity. The typology of PRFSs introduced in this article shows significant variations in the evaluation of the arts within national research funding systems. Those differences—which include varying assessment methodologies and quality criteria—are linked with different relative shortcomings and advantages of the systems, as well as different challenges they create for artists, administrators, research institutions, and artistic disciplines.

6.1 'Artistic research' and 'artistic activity' systems: main challenges

In 'artistic research' systems the arts are classified as a research field and evaluated using the same assessment methods and criteria as other research disciplines. On the one hand, this may have a positive effect on the discipline, as it emphasizes the arts' contribution to knowledge and raises recognition of artists as legitimate research staff (Wilson 2018). On the other hand, because those systems apply the same generic criteria for all disciplines, one of their main challenges is to sustain heterarchy (Stark 2009)—the coexistence of multiple definitions of worth and diverse forms of evaluation. Artistic evaluations within those PRFSs are governed by the concepts of value that originate from research disciplines and artists-academics are expected to adopt those concepts. This includes the necessity of providing 'research statements' and describing artistic outputs in terms of 'research questions', 'research methodology', 'research outcomes', etc. Designers and analysts of 'artistic research' PRFSs emphasize that, without the written supplement, the investigatory dimension of the output may not be clear to evaluators because the research component is often not immediately apparent or deducible from the output itself (Sivertsen 2018; Ysebaert and Martens 2018). In 'artistic research' systems, only the artistic objects that are products of research are eligible for evaluation. However, members of evaluative panels may find it hard to distinguish the outputs of artistic research from the outputs of professional artistic practice, as the difference between practice and research is not manifest in the output but lies in the realization process.

Nevertheless, the requirement of writing research statement raises concerns among artists-academics, who stress that the communication of art experiences in the form of text-based work is often not feasible (McNiff 2007; Blom, Bennett and Wright 2011) and that the construction of a research statement, which includes elements such as research questions, methods, and findings, reflects traditional linear research model not necessarily fitting art. Their claim is that artistic inquiry does not always start with a formulation of a research question and is rarely based on a well-established methodology (Elliott 2011; Trowler 2013). The task of writing the statement is generally believed to be redundant and the PRFSs incorporating such solutions face criticism (Bennett and Franzmann 2013; Harrison and Draper 2014).

The 'artistic activity' PRFSs do not face the challenge of evaluating the research value of artworks because outputs are evaluated in terms of their artistic merit. However, those systems are challenged for other reasons linked with the separation of arts and research embedded in their designs. The inclusion of the arts in *research* funding systems is difficult to justify and the fact that artistic development is supported from research funds can be easily questioned. The classification of artistic disciplines outside the scope of research may suggest that they have academically irrelevant and arguably lower standards than other academic fields. Another problem with those systems is that, while artistic disciplines have their own quality criteria, they share most formal criteria with research disciplines. The formal criteria are, however, based on scientific production norms and styles, which are different from production norms in the arts and therefore inadequate for artistic evaluations (Lewandowska and Kulczycki 2022).[4]

Both types of systems involve substantial investment in data collection. Presentation of outputs plays a crucial role in the evaluation process because panel members assess the submitted material rather than the actual artistic outcomes—that is, they typically do not participate in all artistic events they evaluate. This creates challenges both for PRFS designers who develop output registration systems and for university research and administration staff responsible for outputs' submission. Artistic outputs are much more time-consuming to process and submit than traditional research outputs because the staff needs to provide descriptions, research statements, and diverse verification materials (Barwick and Toltz 2017). Art units tend to submit more items than research units (Barwick and Toltz 2017; Lewandowska and Kulczycki 2022), partly due to different production patterns between arts and research fields and partly because the range of eligible output types is much broader in art than in research disciplines. Although in most PRFSs only a limited number of best outputs need to be submitted, the information about outputs is used by universities for internal staff performance reviews and the research staff members feel pressured to report even marginal artistic outputs as individual 'research' items (Barwick and Toltz 2017; Schippers et al. 2017). This increases the administrative workload affecting art faculties and puts the cost-effectiveness of PRFSs into question.

6.2 'Equal treatment of output types' and 'differential score of output types': main challenges

'Equal treatment' systems are based on the principle that quality is independent of output type. They may be perceived as well suited for the arts as it seems right to assume that all works can be of high or low quality, regardless of their formal classification (a leading role in a performance is not necessarily of better quality than a supporting role, etc). At the same time, those systems are heavily based on peer review which, despite being the most widely accepted form of assessment in academia, has several weaknesses. First, it is subjective and lacks transparency, which may have a negative effect on the perceived fairness of the decision-making process. Second, because of the strong reliance on academic and artistic elites who tend to favour mainstream schools of thought and artistic conventions, it may inhibit artistic or research innovations (see, e.g. Lamont 2009; Hicks 2012; Cattani, Ferriani and Allison 2014; Hug and Ochsner 2022). Peer review also has practical drawbacks, as it has proved to be costly and labour-intensive (Zacharewicz et al. 2019).

'Differential score' systems also use peer review, but the subjectivity of experts is mitigated by the pre-scoring of outputs. The advantage of pre-scoring is higher transparency and increased predictability of evaluation results. However, because scoring is based on outputs' formal properties such as scale (e.g. a theatre piece of 70-min is assigned with more points than a 20-min piece), the point systems reward works that are more labour-intensive but not necessarily artistically better. By confusing productivity with quality, such designed point systems may work against the main goal of PRFSs: to improve the quality of research/artistic activity.

Panel evaluations in 'differential score' systems rely to a greater extent on evidence of artistic quality provided by professional art worlds, such as awards, presentations in prestigious venues, or shortlistings in competitions. Using the traditional forms of artistic outputs' validation in PRFSs may be well received by artists-academics who have more confidence in those mechanisms than in PRFS-panels (Blythe 2018). However, the incorporation of the existing art reviewing systems in PRFSs has its shortcomings. Due to their multiplicity and variety, the validation practices and dissemination channels in the arts are difficult to verify, compare and translate into quality measures. The forms of public assessment are specific to artistic forms of expression and cannot be used as universal indicators in performance-based research funding systems. Although in the sciences the publication channels are uniform and routinely compared using common indicators such as the impact factor (even though, recently, critical voices against simplistic bibliometrics gain power in the sciences as well), in the social sciences and humanities publication channels are more diverse (Kulczycki et al. 2018), which renders evaluative bibliometrics invalid (see for an overview Ochsner 2021). This diversity of dissemination forms is even more pronounced in the arts and, accentuated by the plurality of approaches and aesthetics, it even makes it impractical to construct a single list or rating of venues and modes in which art is made available in the public domain (Lewandowska and Smolarska 2020).

7. Conclusion

This research offers the first overview of performance-based research funding systems that evaluate artistic outputs. Although countries adopt diverse approaches to the evaluation of the arts, we were able to identify three types of systems based on how artistic outputs are classified ('artistic research' vs. 'artistic activity') and evaluated ('equal treatment of output types' vs. 'differential treatment of output types'). This article also identifies 12 quality criteria used by evaluative panels: contribution to knowledge, significance for research, significance for art, originality (intrinsic and extrinsic), rigour, creative context, world-class level, international exposure, peer recognition, output type, and scale of work. The analysis shows that some criteria are used in many countries (contribution to knowledge, significance for research, originality, and peer recognition), whereas some are used only in one or two countries (scale of work and output type).

We found that the criteria correspond to a large degree with those implemented in research assessment regarding the sciences (e.g. Polanyi 1962; Gulbrandsen 2000; Albert, Laberge and McGuire 2012) but also identified criteria specific to arts, such as significance for art and creative context. Moreover, our analysis revealed two types of originality in artistic disciplines. This corresponds well with the findings of Guetzkow, Lamont and Mallard (2004) who identified disciplinary differences in definitions of originality, showing that originality is defined more broadly in the social sciences and humanities than in the natural sciences. Our results add to this broadening of the definition of originality by adding artistic types of originality; in particular, the idea of intrinsic originality (the refinement of work in the course of multiple disseminations) does not seem to exist in research disciplines. It is worth noting that the criteria used in the PRFSs by far do not cover the multitude of quality criteria identified to be relevant for SSH research, such as passion, reflection and criticism, fostering cultural memory, personal voice, etc. (see, e.g. Hug, Ochsner and Daniel 2013; VolkswagenStiftung 2014) and might thus be similarly reductionist for arts research.

The indicators applied in the PRFS under scrutiny are limited in scope even though they are slightly adapted to the arts. As already shown for the humanities, indicators cannot fully reflect research quality even when applying a much more diverse set of indicators (Ochsner, Hug and Daniel 2012). Limiting 'what counts' to a few easily measurable indicators comes with the risk of goal displacement and other negative steering effects (de Rijcke et al. 2016; Xu 2019). The artificial 'scientification' of the arts, that is, putting the (academic) arts in the Procrustean bed of performance indicators developed for the sciences (which do not even work well for the sciences, see Edwards and Roy 2017) comes with the risk of losing the very essence of the art and their societal functions. For example, the imagination and exploration of alternatives or futures as well as ethical considerations, for example, through thought experiments played through in a script for a film or theatre, are different from scientific approaches but are not less rigorous—and offer knowledge generation where scientific approaches fail (e.g. ethical limits). Like science, the arts are about pushing the boundaries and creating knowledge and consciousness. The books of authors like Jules Vernes, George Orwell, or Aldous Huxley or film series like Star Trek or movies like The Matrix amongst others have been important for reflections on data protection, ethical considerations, and technical development, to name a few. Like the sciences, the arts are also about finding 'truth', but it is a slightly different 'mode of truth' or 'veridiction' to use the term by Latour (2012). The integration of the arts into academia offers the opportunity to expand scientific knowledge generation by artistic knowledge generation. If arts are to be academic, they need to integrate and balance out two modes of veridiction. Evaluation procedures therefore must not ask the arts to be scientific but need to not only reflect the dissemination practices in the disciplines but also the knowledge production processes, mode(s) of veridiction and their relevant quality criteria to come to a valid conclusion (see Ochsner et al. 2020).

Future research should investigate this issue by exploring the quality criteria in artistic disciplines. A useful approach in this regard was developed by Ochsner, Hug and Daniel (2013) who derived the assessment criteria from the subjective notions of research quality held by humanities scholars. It seems that a similar approach could be successfully applied in the arts. Furthermore, it would be very helpful to capture qualitatively the experiences and perspectives of panel members who interpret and apply the evaluation criteria. Although our study provides a starting point for a discussion on evaluation in the arts by structuring the criteria included in the evaluation documents, it would be interesting to understand how those generic criteria are interpreted and put into practice in real-life evaluation processes.

By presenting different methods and criteria of evaluation and discussing the challenges related, our work aims to contribute to research as well as evaluation practices. PRFSs are likely to have an increasing impact on art units as their role in the funding of higher art education institutions is rising and institutions with poor results might even be transformed into vocational schools. For art units, which over the last decade have been engaged in research activities and developing research cultures (Harrison and Draper 2014), being a research institution is not only a matter of receiving governmental support but also a question of prestige (Hicks 2012). Although many aspects of PRFSs have been routinely criticized by artists-

academics, art units will be more and more responsive to those systems by adapting to the rules of PRFS evaluation. Therefore, it is of paramount importance that PRFSs take into account the knowledge generation processes in the arts to assure that academia and society can benefit from their specific and unique contribution to knowledge and understanding.

Notes

1. Although in some PRFSs all types of eligible outputs are considered as research outputs, in other PRFSs outputs of artistic activity are not considered as research (see the section 'PRFSs taking into account artistic outputs: an overview'). In this article, we consistently use the term 'artistic outputs' when referring to both types of PRFSs.
2. Although strictly speaking, the Australian Excellence for Research in Australia (ERA) is not a full PRFS given that its results rarely had impact on actual funding decisions, we still included it in the study for the following reasons: First, in 2015, it did have an impact on allocation of funding (Sawczak 2021) and Hicks (2012) classifies it as a PRFS. Second, the other editions of the ERA might not have had an impact on funding but when the exercise was started it was somewhat suggested that it would have. At least, it was said that the results of ERA would inform the governments agendas and that ERA 'allows research managers and investors to identify and reward excellence in research' (Australian Research Council (ARC) n.d., section 'What are some of the benefits of ERA?'). Therefore, Sawczak (2021) concludes that universities only took every edition of the ERA seriously because actual financial consequences were expected. This also means that while it might not have had any actual link to funding, it still did influence research practice as if there was a funding link because the allocation of funding was declared 'not settled' (National Tertiary Education Union n.d., section 'ERA and funding allocation') and only decided after or during the exercise.
3. The Czech Republic evaluates artistic research within PRFS but it has also implemented a separate system for artistic activity, the so-called Register of Artistic Performances (Registr uměleckých výstupů) used for registering and evaluating artistic outcomes of art schools' faculty members. The outputs are assessed by experts and then scored using a mathematical model based on the Saaty-method (see Matějovská and Achten 2020). In 2019, the system was incorporated into the Act on Higher Education Institutions and implemented as a performance-based funding mechanism for the Czech art colleges.
4. For example, in their analysis of the Polish systems, Lewandowska and Kulczycki (2022) have demonstrated that artists-academics tend to produce far more artistic outputs than researchers produce academic publications. Therefore, the number of outputs per researcher (4)—introduced in the Polish PRFS to guarantee that all members of research staff contribute to university research—is irrelevant for artistic disciplines.

Supplementary data

Supplementary data are available at *Research Evaluation Journal* online.

Funding

This work was supported by the National Science Centre, Poland [grant number: 2019/35/D/HS5/00009].

Conflict of interest statement. None declared.

References

Albert, M., Laberge, S., and McGuire, W. (2012) 'Criteria for Assessing Quality in Academic Research: The Views of Biomedical Scientists, Clinical Scientists and Social Scientists', *Higher Education*, 64: 661–76.

Australian Research Council (ARC). (2017) *ERA 2018 Submission Guidelines* (downloaded on 5 February 2021) <https://www.arc.gov.au/excellence-research-australia/key-documents>.

Australian Research Council (ARC). (n.d.) *Excellence in Research for Australia* <https://www.arc.gov.au/excellence-research-australia> accessed 14 April 2022.

Barwick, L., and Toltz, J. (2017) 'Quantifying the Ineffable? The University of Sydney's 2014 Guidelines for Non-Traditional Research Outputs', in R. Burke and A. Onsman (eds.) *Perspectives on Artistic Research in Music*, pp. 67–77. Lanham, Maryland & Oxford: Lexington Books. (*)

Bennett, D., and Franzmann, M. (2013) 'Aristotle and the ERA: Measuring the Immeasurable', *Arts and Humanities in Higher Education*, 12: 367–81. (*)

Bennett, D., Wright, D., and Blom, D. (2009) 'Artist Academics: Performing the Australian Research Agenda', *International Journal of Education & the Arts*, 10: 1–16. (*)

Biggs, M., and Karlsson, H. (2010) 'Evaluating Quality in Artistic Research', in M. Biggs and H. Karlsson (eds.) *The Routledge Companion to Research in the Arts*, pp. 405–24. Routledge: London. (*)

Blom, D., Bennett, D., and Wright, D. (2011) 'How Artists Working in Academia View Artistic Practice as Research: Implications for Tertiary Music Education', *International Journal of Music Education*, 29: 359–73.

Blythe, R. (2018) 'A Validation Model Proposal Based on the Use of Existing Disciplinary Mechanisms as Dashboard Indicators for Validating Research Outputs of Creative or Artistic Works', in B. Van Kerckhoven and W. Ysebaert (eds.) *Evaluating Art and Design Research. Reflections, Evaluation Practices and Research Presentations*, pp. 60–7. Brussels: VUBPress.

Boltanski, L., and Thévenot, L. (2006) *On Justification: Economies of Worth*. Princeton: Princeton University Press.

Brook, L. (2018) 'Evidencing Impact from Art Research: Analysis of Impact Case Studies from the REF 2014', *The Journal of Arts Management, Law, and Society*, 48: 57–69. (*)

Butt, D. (2017) *Artistic Research in the Future Academy*. Intellect Books: Bristol, England. (*)

Cattani, G., Ferriani, S., and Allison, P. D. (2014) 'Insiders, Outsiders, and the Struggle for Consecration in Cultural Fields: A Core-Periphery Perspective', *American Sociological Review*, 79: 258–81.

Croft, J. (2015) 'Composition is Not Research', *Tempo*, 69: 6–11.

Dally, K., Holbrook, A., Graham, A., and Lawry, M. (2004) 'The Processes and Parameters of Fine Art PhD Examination', *International Journal of Educational Research*, 41: 136–62.

de Rijcke, S., Wouters, P. F., Rushforth, A. D., Franssen, T. P., and Hammarfelt, B. (2016) 'Evaluation Practices and Effects of Indicator Use—A Literature Review', *Research Evaluation*, 25: 161–9.

Edwards, M. A., and Roy, S. (2017) 'Academic Research in the 21st Century: Maintaining Scientific Integrity in a Climate of Perverse Incentives and Hypercompetition', *Environmental Engineering Science*, 34: 51–61.

Ek, A.-C., Ideland, M., Jönsson, S., and Malmberg, C. (2013) 'The Tension between Marketisation and Academisation in Higher Education', *Studies in Higher Education*, 38: 1305–18.

Elliott, R. (2011) 'Painting Monkey or Painting Elephant? Some Comments on Measuring Research in the Creative Art' *Australian Universities*', *Review*, 53: 103–9.

Government of the Czech Republic. (2018) *Methodology for Evaluating Research Organisations and Research, Development and Innovation Purpose-tied Aid Programmes (2017+)* (downloaded on 22 January 2021) <https://www.vyzkum.cz/FrontClanek.aspx?idsekce=695512>.

Green, L. (2006) 'Research Outputs in the Creative and Performing Arts: "Australianising" an International Debate', *Media International Australia Incorporating Culture and Policy*, 118: 5–15. (*)

Guetzkow, J., Lamont, M., and Mallard, G. (2004) 'What is Originality in the Humanities and the Social Sciences?', *American Sociological Review*, 69: 190–212.

Gulbrandsen, M. (2000) *Research Quality and Organisational Factors: An Investigation of the Relationship*. Trøndheim: Norwegian University of Science and Technology (NTNU).

Harrison, S., and Draper, P. (2014) 'Evolving an Artistic Research Culture in Music: An Analysis of an Australian Study in an International Context', in S. D. Harrison (ed.) *Research and Research Education in Music Performance and Pedagogy*, pp. 9–25. Dordrecht: Springer Netherlands. (*)

Hellström, T. (2010) 'Evaluation of Artistic Research', *Research Evaluation*, 19: 306–16.

Hicks, D. (2012) 'Performance-Based University Research Funding Systems', *Research Policy*, 41: 251–61.

Hug, S. E., Ochsner, M., and Daniel, H.-D. (2013) 'Criteria for Assessing Research Quality in the Humanities: A Delphi Study among Scholars of English Literature, German Literature and Art History', *Research Evaluation*, 22: 369–83.

Hug, S. E., and Ochsner, M. (2022) 'Do Peers Share the Same Criteria for Assessing Grant Applications?', *Research Evaluation*, 31: 104–17.

Italian National Agency for the Evaluation of Universities and Research Institutes (ANVUR). (2020) *Evaluation of Research Quality 2015-2019 (VQR 2015-2019)* (downloaded on 21 January 2021) <https://www.anvur.it/en/activities/vqr/>.

Johansson, K., and Georgii-Hemming, E. (2021) 'Processes of Academisation in Higher Music Education: The Case of Sweden', *British Journal of Music Education*, 38: 173–86.

Jonkers, K., and Zacharewicz, T. (2017) *'Research Performance Based Funding Systems: A Comparative Assessment'. JRC Science for Policy Report*. Luxembourg: Publications Office of the European Union.

Kessl, F., Lorenz, W., and Otto, H.-U., and White, S. (eds.) (2020) *European Social Work – A Compendium*. Leverkusen: Verlag Barbara Budrich.

Kroll, J., and Webb, J. (2012) 'Policies and Practicalities: Examining the Creative Writing Doctorate', *New Writing*, 9: 166–78.

Kulczycki, E., Engels, T. C. E., Pölönen, J., Bruun, K., Dušková, M., Guns, R., Nowotniak, R., Petr, M., Sivertsen, G., Istenič Starčič, A., and Zuccala, A. (2018) 'Publication Patterns in the Social Sciences and Humanities: Evidence from Eight European Countries', *Scientometrics*, 116: 463–86.

Laiho, A. (2010) 'Academisation of Nursing Education in the Nordic Countries', *Higher Education*, 60: 641–56.

Lamont, M. (2009) *How Professors Think: Inside the Curious World of Academic Judgment*. Cambridge, MA: Harvard University Press.

Latour, B. (2012) *Enquête Sur Les Modes D'existence. Une Anthropologie des Modernes*. Paris: La découverte.

Laudel, G. (2006) 'Conclave in the Tower of Babel: How Peers Review Interdisciplinary Research Proposals', *Research Evaluation*, 15: 57–68.

Lewandowska, K., and Kulczycki, E. (2022) 'Academic Research Evaluation in Artistic Disciplines: The Case of Poland', *Assessment & Evaluation in Higher Education Ahead-of-Print*, 47: 284–96.

Lewandowska, K., and Smolarska, Z. (2020) 'Striving for Consensus: How Panels Evaluate Artistic Productions', *Qualitative Sociology*, 43: 21–42.

Lundén, T., and Sundén, K. (2015) 'Art as Academic Output: Quality Assessment and Open Access Publishing of Artistic Works at the University of Gothenburg', *Art, Libraries Journal*, 40: 26–32.

Matějovská, D., and Achten, H. (2020) 'The Science of Comparing and Scoring Artistic Endeavour at Schools of Higher Education', in L. Werner and D. Koering (eds.) *Anthropologic: Architecture and Fabrication in the Cognitive Age - Proceedings of the 38th eCAADe Conference - Volume 2*, pp. 527–34. Berlin: TU Berlin.

McKee, A. (2020) 'The Criteria Used by Key Decision Makers in Australia to Judge the Academic Quality of NTROs', *Media International Australia*, 177: 165–75. (*)

McNiff, S. (2007) 'Art-Based Research'. In: Knowles J.G., and Cole, A.L. (eds.) *Handbook of the Arts in Qualitative Research*, pp. 29–40. Thousand Oaks: Sage Publications.

Ministry of Education and Science of the Republic of Lithuania (SMM). (2017) *Order on the Approval of the Annual Regulation on Research and Experimental Development and Evaluation of Artistic Activities by Universities and Research Institutes*. <https://e-seimas.lrs.lt/portal/legalAct/lt/TAD/b519a500a93f11e7a65c90dfe4655c64> accessed 11 February 2021.

National Tertiary Education Union (n.d.) *Excellence in Research Australia*. <https://www.nteu.org.au/policy/research/era> accessed 14 April 2022.

New Zealand, Tertiary Education Commission. (2018) *Guidelines for the 2018 Quality Evaluation Assessment Process* (downloaded on 22 January 2021) <https://www.tec.govt.nz/funding/funding-and-performance/funding/fund-finder/performance-based-research-fund/resources-and-publications/#AccordionItem592>.

Nilsson, F., Dunin-Woyseth, H., and Janssens, N. (eds.) (2017) *Perspectives on Research Assessment in Architecture, Music and the Arts: Discussing Doctorateness*. New York: Routledge.

Ochsner, M. (2021) 'Bibliometrics in the Humanities, Arts and Social Sciences', in R. Ball (ed.) *Handbook Bibliometrics*, pp. 117–24. Berlin/Munich/Boston: De Gruyter Saur.

Ochsner, M., Hug, S. E., and Daniel, H.-D. (2012) 'Indicators for Research Quality in the Humanities: Opportunities and Limitations', *Bibliometrie - Praxis und Forschung*, 1: Article 4.

Ochsner, M., Hug, S. E., and Daniel, H.-D. (2013) 'Four Types of Research in the Humanities: Setting the Stage for Research Quality Criteria in the Humanities', *Research Evaluation*, 22: 79–92.

Ochsner, M., Ma, L., Kancewicz-Hoffman, N., Holm, J., Gedutis, A., Šima, K., Hug, S. E., Dewaele, A., and de Jong, S. (2020) *Better Adapted Procedures for Research Evaluation in the SSH: ENRESSH Brief Research Evaluation.* Priziac: ENRESSH. https://dx.doi.org/10.6084/m9.figshare.12049314.

Poland, Ministry of Science and Higher Education. (2019) *Regulation of the Minister of Science and Higher Education of February 22, 2019 on the Evaluation of the Quality of Scientific Activity.* <http://isap.sejm.gov.pl/isap.nsf/DocDetails.xsp?id=WDU20190000392> accessed 13 January 2021.

Poland, Ministry of Science and Higher Education. (2020) *Regulation of the Minister of Science and Higher Education of July 31, 2020 Amending the Regulation on the Evaluation of the Quality of Scientific Activity.* <https://isap.sejm.gov.pl/isap.nsf/DocDetails.xsp?id=WDU20200001352> accessed 13 January 2021.

Polanyi, M. (1962) 'The Republic of Science: Its Political and Economic Theory', *Minerva*, 1: 54–73.

Portugal, Foundation for Science and Technology (FCT). (2018) *Evaluation Guide: R&D Units Evaluation 2017-18*. <https://www.fct.pt/apoios/unidades/avaliacoes/2017/docs/Guiao_de_Avaliacao_Aval_2017.pdf> accessed 12 February 2021.

Research Excellence Framework. (2019a) *REF2021: Guidance on submissions* (downloaded on 4 November 2020) <https://www.ref.ac.uk/publications/>.

Research Excellence Framework (2019b) *REF2021: Panel criteria and working methods* (downloaded on 4 November 2020) <https://www.ref.ac.uk/publications/>.

Riley, H. (2019) 'Aesthetic Cognitivism: Towards a Concise Case for Doctoral Research through Practices in the Visual Arts', *Arts and Humanities in Higher Education*, 18: 430–43.

Rust, C., Mottram, J., and Till, J. (2007) '*AHRC Research Review Practice-Led Research in Art, Design and Architecture*. UK: Arts and Humanities Research Council. (*)

Sawczak, K. (2021) 'Where is the evidence for ERA? Time's up for Australia's research evaluation system', *The Conversation*. <https://theconversation.

com/where-is-the-evidence-for-era-times-up-for-australias-research-evaluation-system-165622> accessed 14 April 2022.

Schippers, H., Tomlinson, V., and Draper, P. (2017) 'Two Decades of Artistic Research: The Antipodal Experience', in J. Impett (ed.) *Artistic Research in Music: Discipline and Resistance: Artists and Reseachers at the Orpheus Institute*, pp. 163–71. Leuven: Leuven University Press. (*)

Sivertsen, G. (2018) 'An Attempt to Create an Indicator for Performance-Based Funding of Artistic Research in Norway', in B. Van Kerckhoven and W. Ysebaert (eds.) *Evaluating Art and Design Research. Reflections, Evaluation Practices and Research Presentations*, pp. 74–8. Brussels: VUB Press. (*)

Slovak Accreditation Agency for Higher Education. (2018) *Methodology for the Evaluation of Standards* <https://saavs.sk/wp-content/uploads/2021/01/Methodology-for-the-Evaluation-of-Standards.pdf> accessed 12 February 2021.

Spain, National Commission for Evaluation of Research Activity (CNEAI). (2020) *Resolution of October 30, 2020, of the National Commission for the Evaluation of Research Activity, By Which the Specific Criteria Approved for Each of the Evaluation Fields Are Published* <https://www.boe.es/diario_boe/txt.php?id=BOE-A-2021-53> accessed 9 February 2021.

Stark, D. (2009) *The Sense of Dissonance: Accounts of Worth in Economic Life*. Princeton: Princeton University Press.

Strand, D. (1998) *Research in the Creative Arts*. Canberra: Department of Employment, Education, Training and Youth Affairs.

Toltz, J. (2020) 'The Value of Australian Artistic Research: How Works Are Assessed, the Place of Peer Evaluation, and the Anxiety of Future Metric Measurement', in *Arts Research Africa Conference 2020 Proceedings*. Johannesburg, South Africa: University of the Witwatersrand Johannesburg. doi 10.17605/OSF.IO/GJDRF. (*)

Trowler, P. (2013) 'Can Approaches to Research in Art and Design Be Beneficially Adapted for Research into Higher Education?', *Higher Education Research & Development*, 32: 56–69.

Vanlee, F., and Ysebaert, W. (2019) 'Disclosing and Evaluating Artistic Research', *Journal of Data and Information Science*, 4: 35–54.

VolkswagenStiftung. (2014) *What is Intellectual Quality in the Humanities*? Hannover: VolkswagenStiftung <https://www.volkswagenstiftung.de/sites/default/files/downloads/Humanities_Quality_Guidelines.pdf>.

Wilson, J. (2011) 'Creative Arts Research: A Long Path to Acceptance', *Australian Universities' Review*, 53: 68–76. (*)

Wilson, J. (2016) 'The White Cube in the Black Box: Assessing Artistic Research Quality in Multidisciplinary Academic Panels', *Assessment & Evaluation in Higher Education*, 41: 1223–36.

Wilson, J. (2018) *Artists in the University: Positioning Artistic Research in Higher Education*. Singapore: Springer. (*)

Wright, D. G., Bennett, D., and Blom, D. (2010) 'The Interface between Arts Practice and Research: Attitudes and Perceptions of Australian Artist-Academics', *Higher Education Research & Development*, 29: 461–73. (*)

Xu, X. (2019) 'Performing under "the Baton of Administrative Power?" Chinese Academics' Responses to Incentives for International Publications', *Research Evaluation*, 29: 87–99.

Ysebaert, W., and Martens, B. (2018). 'The ECOO M-VUB Stakeholder-Driven Evaluation Design for Art and Design Research Outcomes', in B. Van Kerckhoven and W. Ysebaert (eds.) *Evaluating Art and Design Research. Reflections, Evaluation Practices and Research Presentations*, pp. 93–102. Brussels: VUBPress.

Zacharewicz, T., Lepori, B., Reale, E., and Jonkers, K. (2019) 'Performance-Based Research Funding in EU Member States—A Comparative Assessment', *Science and Public Policy*, 46: 105–15.