STUDIES IN HIGHER EDUCATION
https://doi.org/10.1080/03075079.2023.2248174

Routledge
Taylor & Francis Group

Research quality criteria in the Creative Arts

Kamila Lewandowska [a], Michael Ochsner [b] and Emanuel Kulczycki [c]

[a]Science Studies Lab, Centre for European Regional and Local Studies (EUROREG), University of Warsaw, Warsaw, Poland; [b]FORS, Swiss Center of Expertise in the Social Sciences, University of Lausanne, Lausanne, Switzerland; [c]Scholarly Communication Research Group, Adam Mickiewicz University in Poznan, Poznan, Poland

ABSTRACT

This paper investigates research quality criteria in the Creative Arts (CA). The CA has been introduced into the higher education and research sector over the last three decades. It is thus a relatively new research field and there is little empirical knowledge on how outputs in this field should be evaluated. Our study applies a mixed-method approach to assess the relevance of quality criteria used in performance-based research funding systems (PRFSs) in 10 countries. The results of a qualitative analysis of interviews with artists-academics ($N = 67$) and Joint Correspondence Analysis show that when art is evaluated in the context of academic research, both the traditional indicators of artistic quality as well as the cognitive and research-related aspects of the arts are believed to be significant. The JCA analysis also showed that the majority of our respondents found both extrinsic quality criteria (related to reputation and prestige) and intrinsic criteria (related to cognition and development) relevant.

ARTICLE HISTORY
Received 15 March 2023
Accepted 9 August 2023

KEYWORDS
Research evaluation; Creative arts; Quality criteria; Joint correspondence analysis; artistic research

1. Introduction

Research quality criteria are the foundation of university research evaluation systems. However, while those systems are intended to evaluate diverse research fields having different research patterns and quality standards, they tend to adopt universalistic criteria which follow the most dominant research conventions (Gulbrandsen 2000; Hicks 2012) but do not reflect the complex nature of research quality and thus exclude relevant aspects of research (Ochsner 2022). There are several efforts to expand the concept of research quality in evaluations (Andersen 2013; Franssen 2022; Hug, Ochsner, and Daniel 2013; Ochsner, Hug, and Daniel 2013), but they focus on traditional research fields (STEM & SSH) and do not include the field of Creative Arts (CA) which has traditionally been located beyond the context of university research. While the field is sometimes classified as a sub-discipline of the Humanities, it is characterized by substantially different research approaches and production patterns from those specific to the SSH (Lewandowska and Kulczycki 2022). One of the most prominent differences is that outputs in the CA generally do not take the form of scholarly publications but comprise different forms of creative expression such as musical compositions, dance performances, photographic exhibitions, creative written works – the so-called non-traditional research outputs (NTROs, see McKee 2020). Another important difference lies in the fact that the field operates at the intersection of academic research and professional art practice and is a domain of

CONTACT Kamila Lewandowska k.lewandowska11@uw.edu.pl Science Studies Lab, Centre for European Regional and Local Studies (EUROREG), University of Warsaw, Krakowskie Przedmiescie str. 30, 00-927 Warsaw, Poland

practice-based research – a form of research where knowledge is gained by means of practice and the outcomes of that practice (Vear 2022). Therefore, to gain a deeper understanding of how quality is defined and recognized across diverse research fields, current literature on research quality criteria needs to be expanded by developing empirical knowledge on quality evaluations in the CA.

This paper explores research quality criteria in the field of CA. Specifically, we focus on the quality criteria used in the evaluation of the arts within performance-based research funding systems (PRFS) – national systems used to evaluate research outputs and allocate research funding (see, e.g. Hicks 2012). While there is ample research on Art Education and their place and role in Higher Education Institutes (see, e.g. Bequette and Bequette 2012; Eisner 1972; La Porte, Speirs, and Young 2008; Stankiewicz 2000), there are very few studies that have investigated quality criteria used in *research* evaluation systems to assess the CA. The existing research demonstrates that evaluators favour criteria amenable to traditional research conventions and suggest that those criteria may be incompatible with the standards shared by artists-academics (Hellström 2010; McKee 2020). In this study, we adopt a bottom-up approach and explore the notions of quality held by artists-academics, i.e. the evaluated in PRFSs, to better understand how the quality concepts are defined and used within research communities, following the approach suggested by the League of European Research Universities (LERU 2012) to start evaluation procedures on the shop floor to increase the validity of its outcomes (see also Ochsner 2022). The paper starts with identifying the characteristics of research in the Creative Arts. The following sections describe the methods used in this study and discuss the results of our research.

2. Research quality in the creative arts

Thus far, very few studies have empirically investigated research quality criteria in the CA. At the same time, ongoing debates discuss whether research outputs in the field of art should be evaluated using the same methods and quality criteria as other research disciplines. Those debates emanated particularly after the incorporation of art schools into university structures and the harmonization of higher art education with higher education and research systems in the aftermath of the Bologna Process (Kälvemark 2011; Rust, Mottram, and Till 2007; Strand 1998). As a result of those reforms, the CA witnessed the process of academization focused on making higher art education more 'scientific' through linking professional art training to research (Ek et al. 2013). Furthermore, the broadening of the scope of the national research quality assessment exercises to include non-traditional research outputs prompted even more discussions on what constitutes a *research* output and 'to what degree a creative output is the actual research output and not just an addendum, parallel or tangential to the real research?' (Elliott 2011, 2). Based on those debates, we identified six main aspects of research practice in the CA:

2.1. Research problem

One of the most stressed features of the Creative Arts is that the field does not involve the formulation of a 'research question', a 'hypothesis', or a 'research problem' as the first step of the investigation process. Haseman (2006) and Elliott (2011) emphasize that artists do not commence the creative process by setting propositional questions, but instead they tend to 'dive in', to start experimenting right away and see what emerges. Similarly, in Trowler's (2013) study of university-affiliated artists, respondents emphasized the unpredictability and openness of the creation process by saying that the true objective of their work is only to be found 'at the end of the journey' (4). Borgdorff (2011) insists that the requirement that a research study should start with explicit research questions is inadequate for the arts. He states the rationale as follows: 'the artist's tacit understandings and her accumulated experience, expertise and sensitivity in exploring uncharted territory are more crucial in identifying challenges and solutions than an ability to delimit the study and put research questions into words at an early stage.' (56). While illustrating the difference between traditional research and

research in the arts, this citation also points to the similarity: the function seems to be similar, i.e. to explore unchartered territory by developing further existing knowledge or practices, but the form seems to be very different, i.e. the artistic-practical approach precedes the intellectual analysis and putting ideas into words.

2.2. Method

Scholars point out that the choice and application of methods in art is driven by practical concerns rather than guided by methodological orthodoxies. Biggs and Büchler (2007) note that methods in the CA are used in an instrumental way: as practical tools to provide a solution to the problem at hand. Trowler (2013) emphasizes the fact that artists constantly 'reinvent' ways of processing information and 'shift' from one method to another to approach things from a different perspective. Bennett, Wright, and Blom (2009) note that creatives use 'different methods for different aspects of the study' (7) and the creative process is characterized by fluidity and interchangeability of methods instead of methodological transparency. The methodological rigour, therefore, lies not in identifying a method in advance and follow it through but in being fluent in many methodological approaches and apply them rigorously whenever it seems useful.

2.3. References

Andersson (2009) notes that the field of art has its own way of using references, which he describes as follows:

> references selected by artists tend to demonstrate what s/he has been inspired by in terms of art, knowledge, social forces, etc. rather than making explicit an active and clear adherence to or dissociation from something, someone or some theory, which is the purpose of academic references. (4)

By making a reference, an artist positions their work in a certain critical or theoretical context, but they are not obliged to clarify the exact relation between the cited work and their own; the latter is typically done by critics and art historians.

2.4. Originality

Artists emphasize that traditional research and the CA have different definitions of 'originality'. In traditional research, the notion of originality is not easily discernible from the concept of 'innovation'; in the arts, what can be described as 'innovative' is neither unquestionably original nor necessarily has an artistic value. Croft (2015, 8) argues that the projects most easily described as 'innovative' (e.g. a project that involves 'converting arctic ice cap data into sound files to be manipulated in real time in an internet-mediated free-improvisation event') may well lead to music that is conventional, while an original piece for string quartette will be difficult to describe in terms of 'innovation'. Reeves (2016) observes that the term 'innovation' is typically attached to projects that contain technological advancements (e.g. use of new technologies, new instrumental techniques, etc.) but those projects do not necessarily stimulate artistic change.

2.5. Knowledge

The cognitive dimension of art – the fact that art contributes to the production of knowledge – has been highlighted in various studies (e.g. Borgdorff 2011; Butt 2017; Riley 2019). Those studies demonstrate that the type of knowledge typically associated with traditional research – which is linked to the notions of objectivity, reliability, transferability and generalizability (Butt 2017) – is not the only type of knowledge generated within the academic context. Riley (2019) postulates that art is a source of 'non-propositional knowledge', which does not supply generalized laws or

postulate theories but enhances our understanding of the world by helping us realize how we experience it. Therefore, it can include personal, embodied and tacit forms of knowledge that propositional knowledge derived from scientific research does not include. Andersson (2009) points out that knowledge in the arts is communicated in the form of 'concrete oeuvres representing and materialising meaning' (5) and as such it has much more 'tangible' presence than the immaterial knowledge produced by scientists. Due to its concrete representational format, the arts-based knowledge has to be acquired through personal experience rather than by reading an academic paper or a book (Reeves 2016).

2.6. Output

The last point is closely related to the question of output. The field of CA produces mostly non-traditional research outputs (NTROs) which take the form of creative works instead of typical scholarly publications. However, ongoing debates discuss whether the artistic output produced in the academic context should be accompanied by a written supplement describing and evidencing the *research* dimension of the output. Defenders of the supplement emphasize that research elements of the creative process ('purpose', 'methods', 'theoretical context') are not directly deducible from the output itself and, to make them transferable to the research community, they should be clearly explained in a written format (Biggs and Büchler 2008; Ysebaert and Martens 2018). A more liberal approach, and perhaps more popular among artists-academics, is that an artistic output in itself is sufficient for evaluating its contribution and relevance (Lesage 2017). The obligation to attach a 'research statement' to artistic outputs submitted to research evaluation systems (e.g. the UK's REF and the Australian ERA) has therefore come in for a lot of criticism among university-affiliated artists, who emphasize that practice-based research is intrinsically experiential and thus difficult to express in language.

3. Data and methods

This study is part of a research project (*2019/35/D/HS5/00009*) investigating the evaluation of CA within performance-based research evaluation systems (PRFS). A total of 90 interviews with Polish artist-academics were conducted for this project. The first wave included 60 individual in-depth interviews, which examined the effects of the evaluation systems on art schools and their employees; the second wave ($N = 30$; questionnaire-based interviews; labelled as QB [1-30] in Section 4.1) investigated the conceptions of quality in artistic disciplines. In this paper, we analyse qualitative and quantitative data gathered during the second wave as well as interviews from the first wave that included codes relevant to our study, related to the evaluation criteria in artistic disciplines and the relationship between (traditional) research and art ($N = 37$; labelled as IDI [1-37] in Section 4.1). We thus analyse a total of 67 interviews.

The participants of this study were faculty members of 13 Polish art colleges and represented different artistic disciplines, including theatre and film, music, fine arts and conservation. Interviews were conducted online or over the phone (face-to-face interviews were impossible due to the COVID-19 pandemic) between February and November 2021; each interview lasted around 60 min. In-depth interviews conducted during the first wave were semi-structured and based on an open-ended interviewing technique. During the interviews, respondents were asked whether and, if so, how they viewed themselves as 'researchers', if they perceived the effects of their creative work as 'research outputs', and whether they believed the criteria used in the PRFS were adequate quality standards for judging art, and why. Questionnaire-based interviews (the second wave) were conducted using an online questionnaire whose design drew on earlier stages of the current research project. The questionnaire listed 12 quality criteria used in 10 PRFSs (in Australia, the Czech Republic, Italy, Lithuania, New Zealand, Poland, Portugal, Slovakia, Spain, and the United Kingdom) to evaluate artistic outputs (Table 1). The criteria were identified through an analysis of

Table 1. Relevance of research quality criteria in the CA.

Criteria	Relevant	Somehow relevant	Irrelevant	M (SD)
Significance for research	36.67%	36.67%	26.66%	3.07(1.17)
Significance for artistic development	80.00%	20.00%	0.00%	4.30(0.79)
Contribution to knowledge/understanding	76.67%	16.67%	6.67%	4.00(0.87)
Rigour (methodological; artwork as a results of a systematic research process)	16.67%	30.00%	53.33%	2.33(1.09)
Creative or intellectual context (references)	76.67%	20.00%	3.33%	4.23(0.90)
Originality: extrinsic (a new contribution to art or research)	60%	36.67%	3.33%	3.83(0.87)
Originality: first presentation	N/A	N/A	N/A	N/A
World-class level (the output ranks with the best within its discipline)	66.67%	23.33%	10%	3.83(1.05)
International exposure (dissemination)	60.00%	23.33%	16.66%	3.57(1.04)
Peer recognition (e.g. awards, festival invitations)	76.66%	13.33%	10.00%	3.97(1.03)
Scale of work (e.g. duration of a play, cast size)	10%	26.67%	63.34%	2.13(1.11)
Output type (e.g. a concert, a solo exhibition)	43.34%	33.33%	23.33%	3.17(1.09)

evaluation regulations and guidelines published by the national evaluating agencies (see for more details Lewandowska et al. 2023). Evaluation panels are required to follow those guidelines and apply the quality criteria listed in them in the evaluation process. However, the guidelines are very general, and evaluators are given no specific instructions on how to interpret and operationalize the criteria.

During questionnaire-based interviews, the respondents were asked to complete the questionnaire by rating the importance of each criterion for judging the quality of artistic outputs on a 5-point Likert-type scale. In addition, respondents were asked to explain how they understood the criteria and to justify their ratings (they were asked questions such as: 'Why do you think 'peer recognition' is *very relevant*?' or 'Why have you scored 'significance for artistic development' more highly than 'significance for research''?). This mixed-method approach (a questionnaire and an interview) was applied because the criteria drawn from evaluation documents were generic concepts and it was necessary to understand the subjective meanings our respondents attached to them to be able to draw meaningful conclusions. Table 1 lists the criteria and presents descriptive statistics.

All interviews were tape/video-recorded and transcribed. Interviews were carried out in Polish and transcribed as such before being translated into English. Qualitative data were analysed with the assistance of MAXQDA. Transcripts were coded line-by-line and the codes were developed inductively. Quantitative data gathered through the online survey were recorded in Excel and analysed using R and Stata 16. Given the small sample size of 30 questionnaires, we opted for an explorative analysis using robustness tests to examine the influence of outliers on the results. Specifically, we applied Joint Correspondence Analysis (JCA), a more efficient variant of Multiple Correspondence Analysis (Camiz and Gomes 2013; Greenacre 2007). Correspondence Analyses can be considered the equivalent to Principal Component Analysis for categorical variables (Fithian and Josse 2017). The advantage of Correspondence Analysis is that no statistical model is involved and, hence, no assumptions regarding distributions, relationships, etc. need to be made (de Leeuw and van der Heiden 1988; Di Franco 2015; Fithian and Josse 2017). Furthermore, empirical tests have shown that it is especially robust for small-n high-dimensional analyses (Fithian and Josse 2017) even though from theoretical point of view it is often suggested that bigger samples are used for stability reasons (Di Franco 2015). A big advantage is that Correspondence Analyses focuses on displaying relationships between variables and observations (Di Franco 2015).

We apply JCA in the Benzecri tradition, i.e. as merely a graphical representation of the data at hand, i.e. a simplification of a complex table. For the JCA to converge, we had to simplify the data structure by recoding the 5-point Likert scale into dummies (Di Franco 2015). We chose to focus on the respondents who find the criterion relevant or very relevant (we coded answer categories 4 and 5 into 1 and 1–3 into 0). This choice was made both for interpretational reasons (see below), but also given the distributions as some criteria were only rated as irrelevant (1 or 2) by only very few respondents.

Due to the small sample size, we conducted robustness tests. First, we rerun the JCA taking out two cases. Cases were put into pairs randomly. Second, we rerun the JCA 18 times by randomly dropping two cases with replacement (i.e. the same case could be dropped in two different renditions). Finally, we also conducted a Principal Component Analysis with oblimin rotation to see whether the choice to use dummies makes a difference. Note, however, that the PCA comes with several distributional assumptions that cannot be upheld, i.e. our rating scale is non-equidistant ordinal while the PCA assumes a linear (and unbound interval) scale; also, PCA is at least as sensitive to small sample sizes, even though it is often used in high-dimensional analyses, i.e. where more variables than observations are considered (but usually with sample sizes clearly higher than 30).

4. Results

4.1. Relevance of the criteria

Drawing on the analysis of interviews and using the percentage distribution of survey responses, we identified criteria that were assessed as: ***important*** *(the most relevant criteria*, i.e. more than 75% of respondents assessed the criterion as 'very relevant' or 'relevant'), ***adequate*** *(relevant criteria*, i.e. 50–75% of respondents indicated it was very relevant or relevant), ***contentious*** *(the least relevant criteria*, i.e. less than 50% of respondents indicated relevance). In Table 1, responses on the 5-point Likert scale were regrouped into 3 categories: Relevant (categories 4 and 5 on the Likert scale), Somehow relevant (category 3, which was generally chosen by respondents as an 'OK, but … ' middle option), Irrelevant (categories 1 and 2).

One criterion (*Originality: first presentation*) was excluded from the further analysis based on the analysis of interviews. This criterion is used in PRFSs to indicate that only the first presentation of an output (a premiere, first performance, etc.) is eligible for evaluation and repeated exhibitions or events cannot be submitted as multiple outputs. Respondents who interpreted the criterion in this fashion generally found it irrelevant. However, other respondents understood this criterion as internal originality ('an artistic work evolves over time and each presentation has a new component') and found it very relevant. Due to these contradictory interpretations, it was impossible to draw meaningful conclusion from the quantitative data for this criterion.

Important criteria included *Significance for artistic development*, *Contribution to knowledge/ understanding*, *Creative or intellectual context*, *Peer recognition*.

Significance for artistic development was the most approved criterion. In explaining why they have scored this criterion highly, artists commented that high-quality outputs 'move things forward' (QB_14), 'set new trends' (IDI_26) and 'say something new aesthetically' (QB_4). A slightly different view was presented by respondents who noted that art does not follow a linear process of development (IDI_32) and their impact is clear only after a considerable amount of time (QB_10; QB_12) and therefore this criterion would be difficult to operationalize within research funding systems. In making this point, some made comparisons with science by saying that progress in the scientific fields is much faster and much more 'obvious' than in the arts where 'we only know after a decade or so that it was significant and constitutive for art' (QB_10).

Another important criterion was *Contribution to knowledge/ understanding*. The interviewees emphasized the cognitive aspects of art by saying that art 'helps understand things' (QB_9), 'gives us insights into the nature of existence' (QB_12), 'inclines people to reflect' about the current issues (An_6). Art contributes to knowledge/ understanding by offering ways of learning that are parallel to scientific cognition; as one respondent puts is, art is 'one of the ways of gaining knowledge about the world, and scientific research is another way' (IDI_19). This perspective is similar to Riley's (2019) concept of *aesthetic cognitivism*, which implies that, while art may not be a source of knowledge in a conventional sense (knowledge that is obtained through scientific research), it is a powerful means of deepening our understanding and, in this respect, it provides access to an alternative ('non-propositional') type of knowledge.

Creative or intellectual context is also among the most relevant criteria. Respondents would expect a high-quality work of art to be located in a 'wider intellectual context' (QB_12), which means that is draws upon ('cites') other significant artistic or intellectual achievements. Similarly to McKee's study of the evaluators in the Australian ERA (2020), our respondents thought it was relevant to judge artistic outputs in terms of how well they engage with the current state of knowledge and development in the CA. At the same time, some of our respondents took a slightly different route and emphasized the *social* context, pointing out that good art 'does not refer only to itself' (QB_9) but reflects upon social issues and drives change.

Peer recognition was also found to be important. Our respondents highlighted the role of peers as the 'primary assessors' in artistic disciplines (QB_16) and insisted that art communities 'have the best available expertise' for making artistic judgements (IDI_5). Peer assessment was described using words such as 'reliable' (QB_9), '(relatively) fair' (QB_12), even 'objective' (IDI_36) and respondents felt that the traditional indicators of artistic prestige, such as art awards or inclusion into prestigious art collections, were a trustworthy evidence of quality (QB_13). While a few respondents called attention to the issues of subjectivity and bias in peer evaluation, there seemed to be a general agreement that decision-making in artistic disciplines should be conducted through a community-led peer review – which means that evaluations are performed by members of discipline-based art communities. The arts have extensive peer review traditions and well-developed systems of validating artworks, and previous studies have shown that artists tend to trust those systems more than they trust research evaluation systems (Blythe 2018).

Adequate criteria were the following: *Originality: extrinsic*, *World-class level*, *International exposure*.

Originality: extrinsic was assessed as an adequate criterion. Similarly to Guetzkow, Lamont, and Mallard (2004), in our study, respondents referred to different types of originality, which included 'new interpretation' (e.g. a novel staging of a classic play), 'new approach' (e.g. a work that 'offers a new perspective on art'), 'new topic' (e.g. a work 'says something new about the world'), 'new aesthetics' (a work 'goes beyond aesthetic canons' (IDI_35)), 'technical/technological' originality (an artist 'advances graphic techniques' (IDI_37)).

According to our respondents, originality is relevant but not the most important as it is 'not the main objective of art'. Firstly, not every work of art has to be original to be worthy. In making this point, artists made comments along the lines of: 'Vermeer's paintings were not particularly original even in his times, but they were perfectly composed, remarkable in terms of light, mood, etc.' (QB_3) or 'An interpretation of Chopin's piano concerto does not have to be original to be moving' (IDI_29). Secondly, doing something new is not necessarily the same as doing something worthwhile. Some respondents noted that extreme innovations in instrumental technique can be aesthetically regressive. Thirdly, contemporary art theories have undermined the cult of originality in the arts, by arguing that it is an 'overrated and out-of-date' category (IDI_9). In line with this view, respondents emphasized that the notion of originality is simply not possible anymore because 'artists always borrow from other artists' (QB_9). Another respondent insisted that 'an artistic expression that is entirely unique and has not received from other works of art in any way, does not exist' (QB_5).

Adequate criteria also included *World-class level* and *International exposure*. The difference between those two criteria was explained to respondents by the interviewer: *World-class level* is a quality benchmark that *does not* refer to the place of dissemination (there is no assumption of any necessary international exposure), while *International exposure* specifically addresses cross-national distribution. However, this distinction was impractical for the artists who felt that it would be very difficult to demonstrate a world-class quality of an output that has not been presented in an international arena. They noted that *World-class level* would be difficult to judge without presenting evidence of international acclaim, such as art prizes, exhibitions at the world's best galleries, invitations to the top festivals, etc. Respondents thus found those criteria to be very closely related.

International exposure is a relevant criterion because art distributed cross-nationally is validated by larger and more diverse audiences (QB_11) and goes through more competitive peer-reviewing

processes (QB_1). However, respondents stressed that this criterion would make sense in relation to some genres but not others – for example, international exposure of theatre productions is difficult due to language barriers. It was also repeatedly highlighted that international dissemination is largely dependent on access to resources (travel funding, translation grants, etc.), which are typically accumulated by a relatively small group of the most reputable artists.

Contentious criteria included: *Significance for research*, *Rigour*, *Scale of work*, *Output type*.

Significance for research is among the least relevant criteria; however, the frequency percentage distribution and the relatively high *standard deviation* (a proxy of disagreement among respondents) suggest that this criterion is more non-consensual (contentious) than simply irrelevant. In addition, the analysis of qualitative data showed that this criterion is understood and perceived differently by different groups of respondents. Specifically, we identified four different approaches to the evaluation of this criterion's relevance among artists-academics:

The first approach, which we call the *Isolationist-negative*[1] perspective was represented by artists-academics who emphasized the opposition between art and 'scientific research'. Respondents stressed that artistic creation and scientific research are two independent and substantially different forms of activity. In making this point, they made comments that fitted into the widespread popular understanding of art as the opposite of science: 'We are artists. We make art, we don't do science' (IDI_20), 'Science appeals to reason, art appeals to emotions' (QB_7), 'The word "scientific" makes me think of something structured, systematic. And theatre is a place where … it's like a madhouse' (QB_12). Those respondents spoke critically about the inclusion of the CA into the academic systems and the necessity of describing artistic outputs using the language of academic research. They called such practices 'cognitive falsehood' (IDI_35), 'a sham' (IDI_20), 'nonsense' (QB_7), or 'pseudo-science' (QB_15).

The somewhat different, *Isolationist-positive* approach was represented by interviewees who focused on the areas of intersection between art and academic research. The most obvious example of those intersections are disciplines focused on the analysis of art (art history, theatre studies, musicology, etc.) but respondents also referred to more interdisciplinary fields combining, for example, art and medical sciences (art therapy), art, engineering and technology (architecture, design), art and chemistry (art conservation). This group of respondents also highlighted the meta-critical role of art; namely, that art contributes to academic research by offering important reflections on ethical issues, technological development, climate change, etc. Therefore, while art and research are separate forms of activity, some forms of art can be significant for some forms of research.

The third and fourth position are labelled as the *Situated* approach. From the *Situated* perspective, art and research are not exclusive, but overlapping domains. However, respondents seemed to perceive this overlap in two different ways. The *Situated-hard* approach is represented by respondents who spoke of artistic process *as a form of* research. As one respondent put it: 'we constantly search, reflect and ask questions about the practical implications of our artistic choices' (IDI_1). For those artists, research – in the form of self-inquiry, experimentation, and critical reflection – is an essential part of every artistic process.

A different group, which we label as the *Situated-soft* included respondents who believed that *some* forms of art *can be* research. This group focused on projects whose primary aim is to research a particular subject using art-based methods (e.g. 'Kevin Lee's video essays, which explore films through an audio-visual analysis' IDI_19), or highlighted specific components of artworks that align with scientific method (e.g. experimentation with new technological solutions (IDI_15)).

Methodological rigour proved to be a more problematic criterion. Respondents associated this criterion with a mechanistic and standardized approach to research, which, in their view, does not correspond with how art is typically made. They stressed that, unlike traditional academic research, artistic process is neither linear nor systematic; methodological aspects can only be identified retrospectively, *ex post*. As one respondent put it:

> Rigour in science gives structure to the [scientific] process. If you follow the rules, you're safe. Here [in art] rigour is a different thing … Methodological tools are applied depending on how a specific problem unfolds or how it escapes us. And, what is perhaps most important, sometimes we abandon all methods and spontaneously take a creative risk. And I think this is the most appropriate way to make art: to take a risk and follow your creative impulse. (QB_16)

The respondent highlights the difference between art method and scientific method: in art, methods are chosen, re-evaluated and modified during the creation process, in response to the problem at hand, while in science the research process is fixed. This idea of scientific method as a procedure (rather than a strategy) clearly influences the way artists-academics understand *rigour* and assess their relevance as a quality criterion in the arts. The description of artistic process suggests that art is, in fact, rigorous, as it requires consistency: a particular method is chosen (or abandoned) because of its (in)appropriateness to deliver the research solution.

Scale of work was the least approved criterion. Respondents insisted that scale (e.g. duration of a performance, cast size, physical size of a sculpture) is not related to artistic quality. Those respondents made comments such as: 'the size does not determine the value. What I mean is … it would be difficult to judge quality of a painting by whether it is one meter by one meter or 20 centimetres by 20 centimetres' (QB_9) or 'a brilliant performance made by two people can be much more significant than a large-scale performance by a symphony orchestra' (An_10). Another artist noted that this criterion is particularly irrelevant for the evaluation of contemporary art which values ideas above the formal or visual components of artworks: 'you wouldn't judge a piece of art simply based on how it looks like, how large it is, or what kind of fabric has been used, right?' (IDI_36).

At the same time, respondents were more favourable towards using the *Output type* criterion, which indicates that some types of outputs (e.g. a music recital) would be scored more highly than others (e.g. a supporting part in a ballet performance). In contrast to *Scale of work*, which was associated by respondents with measurable and formal aspects of outputs, this criterion was related to the assessment of the artist's contribution to an artistic production. Respondents who assigned importance to this criterion, insisted that 'a leading part and a background role in a movie are not of the same importance' (QB_1) or 'there is a difference between a solo exhibition, where all responsibility falls on the main author, and a contribution to a collaborative group show' (QB_9). Therefore, artists-academics seem to disapprove the criterion of scale when it is understood in a merely formal sense (e.g. long-lasting performances are scored higher than short performances), but when the criterion addresses the amount of creative input – for example, a portfolio of five photographs is scored more highly than each of those photographs submitted as an individual output (see McKee 2020) – then the criterion becomes more acceptable.

4.2. Patterns among artists-academics

To identify patterns how artists-academics see the relevance of quality criteria, we conducted a Joint Correspondence Analysis. We excluded the *Scale of work*, however, as it was too strongly influencing the solution as only three respondents qualified this criterion as relevant. While the solutions with or without *Scale of work* were almost identical when it comes to interpretation, in the robustness tests, the position of this criterion changed considerably depending on cases included or excluded and with it also the position of the other items changed somewhat. We therefore focused only on 10 criteria for the JCA. Given the high interpretability of the graph and the low sample size, we only interpret a two-dimensional map (see Figure 1); a robustness test applying a principal component analysis also suggested to retain two factors. The first dimension clearly separates the relevance or non-relevance of the criteria. This dimension thus reflects the level of agreement with the PRFSs' criteria of quality: the more to the right, the more criteria used to assess the CA within the PRFSs are accepted (perceived as relevant) by the respondents. On the second dimension, items load high that reflect the external attribution of quality (the criteria related to reputation and prestige), for example *World-class level*, *International exposure* or *Peer recognition*, while criteria reflecting

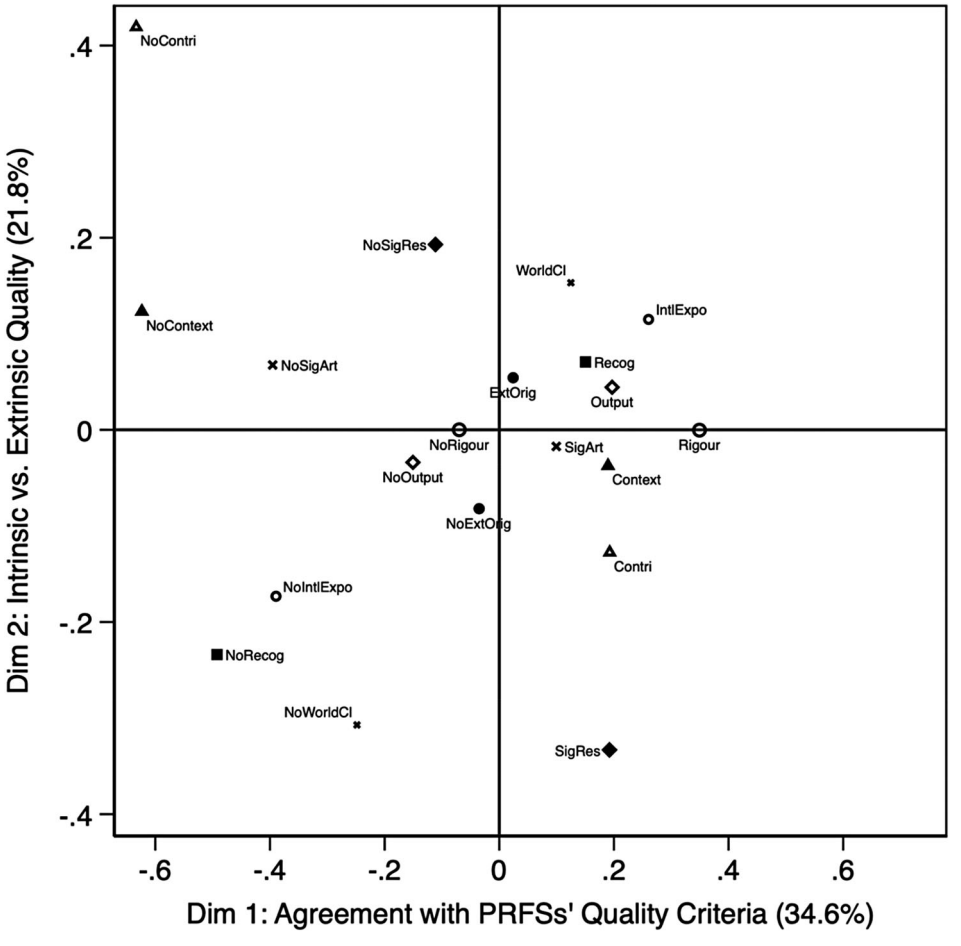

Figure 1. Criteria in the two-dimensional space from the Joint Correspondence Analysis.

intrinsic quality (related to the aspects of cognition and development), such as *Significance for research* and *Contribution to knowledge/understanding,* score low. This second dimension therefore reflects the intrinsic vs. extrinsic quality.

We thus see that there is a difference between the level of agreement of the PFRSs' quality criteria and the attribution of quality (intrinsic vs. extrinsic). But we can also see from Figure 1 that the two dimensions are correlated. Figure 2 shows the same map but with the respondents plotted onto it. Three cases stand out, i.e. 4, 6, and 9. They do not find many PRFSs' criteria relevant (zero to two criteria were rated as relevant). Respondent 4 finds only *World-class level* relevant, Respondent 9 only *Significance for artistic development* and *Originality: extrinsic*. This group represents artists-academics who do not believe that the quality criteria used within the PRFSs are useful to evaluate artistic research.

The second group of respondents, i.e. respondents 10, 13, 24, 28, stands out for finding extrinsic quality criteria particularly relevant to evaluate artistic research. They all find *World-class level*, *Peer recognition*, *Significance for artistic development*, *Originality: extrinsic*, and (all but respondent 10) *International exposure* relevant, but not *Significance for research*, *Contribution to knowledge/understanding*, and *Rigour*. The respondents rated five to seven criteria as relevant. This group includes artists-academics who find traditional forms of artistic validation relevant, but do not find research-related criteria adequate for making assessments in the CA. Thus, this group's attitude

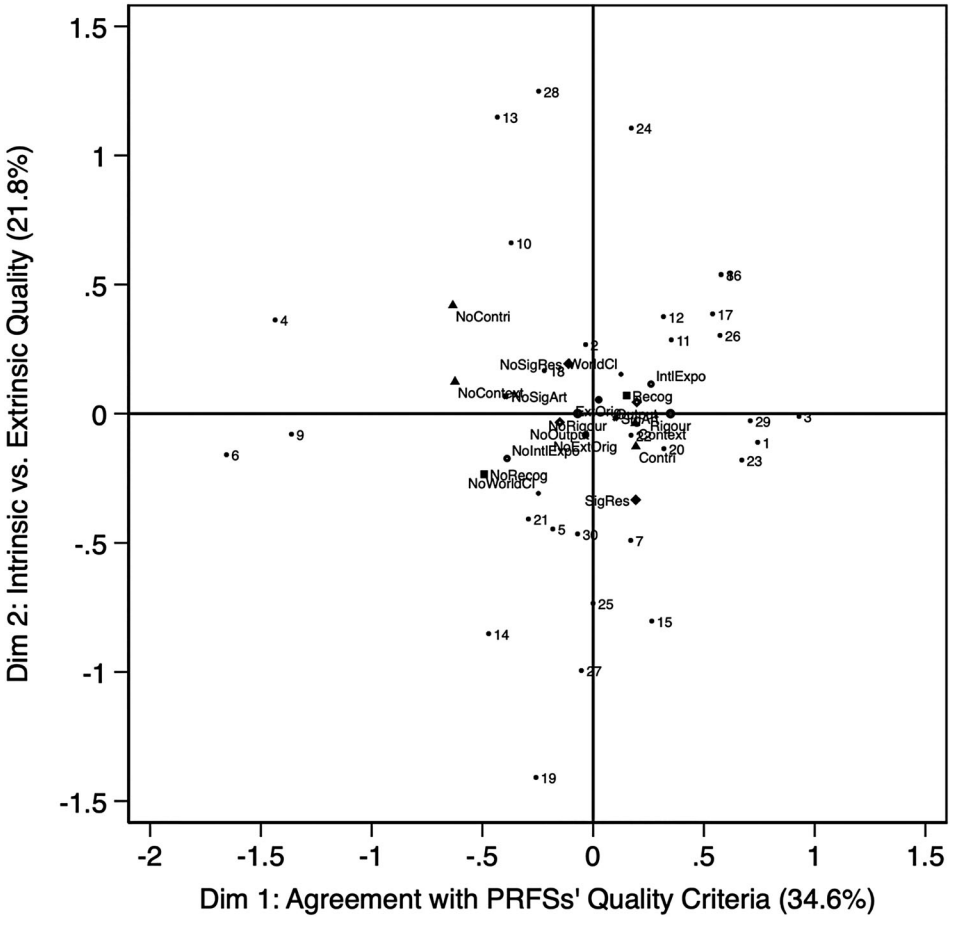

Figure 2. Biplot of the Joint Correspondent Analysis (both observations and criteria in the same geometrical space).

towards the evaluation of the CA seems to be most compatible with the *Isolationist-negative* approach, identified earlier through the interviews.

A third group consisting of respondents 8, 11, 12, 16, 17, 26 finds a mix of intrinsic and extrinsic criteria relevant (i.e. *Contribution to knowledge/understanding*, *Creative or intellectual context*, *World-class level*, *Peer recognition*, *International exposure)* but not *Significance for research* and *Scale of work*. All but respondent 12 find also *Significance for artistic development* relevant. Respondents in this group rated six to eight criteria as relevant. This group represents artists-academics who find criteria related to reputation and prestige (extrinsic quality) relevant but also some criteria related to the cognitive aspects of art; however, not necessarily the ones associated with traditional research, such as *Rigour*. This group represents a more nuanced ('soft') *Isolationist* approach as it finds the cognitive aspects of art important.

The fourth group lies at the right-hand side on the *x*-axis and consists of respondents 1, 3, 23, 29. These respondents find most criteria relevant (nine to eleven criteria rated as relevant). Similarly to group three, the respondents find *Contribution to knowledge/understanding*, *Creative or intellectual context*, *World-class level*, *Peer recognition*, *International exposure*, *Significance for artistic development* relevant; but they also find *Significance for research* relevant as well as *Originality: extrinsic* (not respondent 23, however) and *Output type* (not respondent 1). This group finds most criteria relevant, both reputation-based as well as research-based, and may correspond both with the *Isolationist-positive* or *Situated-hard (art is research)* perspective.

The fifth group is situated at the bottom in the middle. Respondents 14, 15, 19, 25 and 27 belong to this group who find five to seven criteria relevant. They prefer the intrinsic criteria, such as *Significance for research*, *Significance for artistic development*, *Contribution to knowledge/understanding* and (except for respondent 14) *Creative or intellectual context*, while they do not find *Rigour* relevant neither as *World-class level* (except for respondent 15) and *International exposure* (except for respondent 25). This group thus represents artists-academics whose approach to the evaluation of the CA is most compatible with the *Situated-soft (artistic research)* perspective.

Finally, some respondents are located in the middle of the graph and have thus a less pronounced characteristic. They find five to seven criteria relevant. All of them find *Contribution to knowledge/understanding* relevant and all but respondent 30 *Creative or intellectual context*, while they all deem *Scale* irrelevant and all but respondent 30 *Rigour.* Only half of them find *World class level*, *International exposure* or *Peer recognition* relevant and nobody finds more than two of them relevant. They are thus artists-academics who have a tendency to find intrinsic criteria relevant and some of the reputation-oriented criteria.

5. Conclusion

In this article, we have attempted to explore the acceptance of the quality criteria used to evaluate the CA within the PRFSs. With increasing pressure in the science and higher education sectors, this field is also subject to various pressures of accountability and quantification. Nevertheless, as our comprehensive analysis of the evaluation of the arts within performance-based research funding systems (Lewandowska et al. 2023) has shown, the criteria for evaluating quality in the CA can sometimes be both strongly vague and undefined, as well as – at the same time – very precisely quantified, as in the case of STEM or SSH. Therefore, we undertook a study of how the quality criteria used in the various PRFSs in the CA field are understood by academics, which criteria applicable to research evaluation they consider relevant, and which they consider redundant and mismatched.

We discovered that four criteria were considered important by a large majority: *Significance for artistic development*, *Contribution to knowledge/understanding*, *Creative or intellectual context*, *Peer recognition*. This result suggests that when art is evaluated in the context of academic research, both the traditional indicators of artistic quality (art prizes, etc.) and the cognitive aspects of art are believed to be significant. The ratings of the quality criteria also showed that the majority of our respondents found both extrinsic quality criteria (related to reputation and prestige) and intrinsic criteria (related to cognition and development) relevant. Our study thus demonstrates that the widespread popular notion of art as merely affective and noncognitive (Lesage 2017) does not correspond with how the majority of our respondents view the field they practice. Furthermore, the qualitative analysis suggested a diversity of approaches concerning the relationship between art and research: the Isolationist-negative ('art is art, research is research'), Isolationist-positive ('art and research influence each other'), the Situated-hard ('art is research'), the Situated-soft ('art can be research'), and we re-identified those approaches using the JCA method. This diversity of perspectives shows that research evaluation systems that situate art in opposition to research or develop art-specific criteria instead of broadening the definition of research to include arts-based approaches (Biggs and Büchler 2007) may not reflect the complex nature of research in the CA.

Three criteria were assessed as adequate: *Originality: extrinsic*, *World-class level*, *International exposure*. Similarly to the SSH (Guetzkow, Lamont, and Mallard 2004), the Creative Arts have different types of originality but also include types that are specific to the arts (*originality: first presentation*). Moreover, like in the humanities, in the Creative Arts originality is perceived as *a*, rather than *the* criterion of research quality (VolkswagenStiftung 2014), which means that it is relevant but some other aspects of a creative work related to its aesthetic, cognitive, art-historical, or interpretation-centred value may be at least as important. Another characteristic shared with the humanities is that, because many artistic productions are rooted in particular cultural traditions or expressed using national or regional languages (Archambault et al. 2006; Hellqvist 2010), the

possibilities of their international exposure are limited. This issue is particularly relevant in more humanities-related art fields where text plays a significant role (e.g. theatre, rap music) but may have less effect on, for example, visual arts or classical music.

Four criteria were assessed as contentious: *Significance for research*, *Rigour*, *Scale of work*, *Output type*. Our results suggest that certain research-defining criteria, such as method or rigour, should be revisited and refined to include practices specific to the CA rather than excluded from the field. As our results demonstrate, those criteria are frequently reduced to the narrow definitions of 'scientificity', not only by the decision-makers designing the evaluation systems but also by the arts communities themselves. Biggs and Büchler (2007) stress that those communities should aim for broadening the definition of research used in the PRFSs rather than insist on implementing discipline-specific criteria. We believe that future studies could focus less on emphasizing the differences between the CA and other research disciplines and more on developing the notions of research quality that reflect the characteristics of research in the CA but also preserve the importance of methodological and intellectual rigour.

Empirical studies investigating research quality criteria are often conducted in particular national contexts (e.g. Guetzkow, Lamont, and Mallard 2004; Hellström 2010; McKee 2020; Ochsner, Hug, and Daniel 2013) which may limit the generalizability of findings. Our study used criteria from 10 national research evaluation systems but we surveyed only researchers in Poland. This is a possible limitation of this study; however, the fact that many aspects of research in the CA raised by our participants were already described in previous work focused on different country contexts (e.g. Biggs and Büchler 2007; Borgdorff 2011; McKee 2020) indicates that our findings may have validity also outside the higher education and research sector in Poland. Regardless, future research could continue to explore quality criteria in the CA using larger and international samples.

Note

1. The concept of Isolationist and Situated positions in the Creative Arts was originally proposed by Biggs and Büchler (2008). The Isolationist Position states that the CA is significantly different from and incomparable with other research fields, and the Situated Position indicates the opposite by emphasizing the commonalities and comparability between the CA and academic research. Based on the results from our study, we modified this binary classification by adding further distinctions ("positive", "negative", "hard", "soft").

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

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

ORCID

Kamila Lewandowska http://orcid.org/0000-0002-2189-239X
Michael Ochsner http://orcid.org/0000-0002-1758-9590
Emanuel Kulczycki http://orcid.org/0000-0001-6530-3609

References

Andersen, J. P. 2013. *Conceptualising Research Quality in Medicine for Evaluative Bibliometrics*. Det Humanistiske Fakultet, Kobenhavns Universitet.

Andersson, E. 2009. "Fine Science and Social Arts – on Common Grounds and Necessary Boundaries of Two Ways to Produce Meaning." *Art and Research* 2 (2).

Archambault, É, Vignola-Gagné, É. Côté, G. V. Larivière, and Y. Gingrasb. 2006. "Benchmarking Scientific Output in the Social Sciences and Humanities: The Limits of Existing Databases." *Scientometrics* 68 (3): 329–42. https://doi.org/10.1007/s11192-006-0115-z.

Bennett, D., D. Wright, and D. Blom. 2009. "Artist Academics: Performing the Australian Research Agenda." *International Journal of Education & the Arts* 10 (17): 1–16.

Bequette, J. W., and M. B. Bequette. 2012. "A Place for ABT and DESIGN Education in the STEM Conversation." *Art Education* 65 (2): 40–7. https://doi.org/10.1080/00043125.2012.11519167.

Biggs, M., and D. Büchler. 2007. "Rigor and Practice-based Research." *Design Issues* 23 (3): 62–9. https://doi.org/10.1162/desi.2007.23.3.62.

Biggs, M., and D. Büchler. 2008. "Eight Criteria for Practice-based Research in the Creative and Cultural Industries." *Art, Design & Communication in Higher Education* 7 (1): 5–18. https://doi.org/10.1386/adch.7.1.5_1.

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 *Evaluating art and Design Research. Reflections, Evaluation Practices and Research Presentations*, edited by B. Van Kerckhoven and W. Ysebaert, 1st ed., 60–7. Brussels: VUBPress.

Borgdorff, H. 2011. "The Production of Knowledge in Artistic Research." In *The Routledge Companion to Research in the Arts*, edited by M. Biggs and H. Karlsson, 1st ed., 44–63. London: Routledge.

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

Camiz, S., and G. C. Gomes. 2013. "Joint Correspondence Analysis Versus Multiple Correspondence Analysis: A Solution to an Undetected Problem." In *Classification and Data Mining*, edited by A. Guisti, G. Ritter, and M. Vichi, 11–8. Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-642-28894-4_2.

Croft, J. 2015. "Composition is Not Research." *Tempo* 69 (272): 6–11. https://doi.org/10.1017/S0040298214000989.

de Leeuw, J., and P. G. M. van der Heiden. 1988. "Correspondence Analysis of Incomplete Contingency Tables." *Psychometrika* 53 (2): 223–33. https://doi.org/10.1007/BF02294134.

Di Franco, G. 2015. "Multiple Correspondence Analysis: One Only or Several Techniques?" *Quality and Quantity* 50 (3): 1299–315. https://doi.org/10.1007/s11135-015-0206-0.

Eisner, E. W. 1972. *Educating Artistic Vision*. New York: Macmillan.

Ek, A.-Ch., M. Ideland, S. Jönsson, and C. Malmberg. 2013. "The Tension Between Marketisation and Academisation in Higher Education." *Studies in Higher Education* 38 (9): 1305–18. https://doi.org/10.1080/03075079.2011.619656.

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

Fithian, W., and J. Josse. 2017. "Multiple Correspondence Analysis and the Multilogit Bilinear Model." *Journal of Multivariate Analysis* 157: 87–102. https://doi.org/10.1016/j.jmva.2017.02.009.

Franssen, T. 2022. "Enriching Research Quality: A Proposition for Stakeholder Heterogeneity." *Research Evaluation, Advance Access*, https://doi.org/10.1093/reseval/rvac012.

Greenacre, M. 2007. *Correspondence Analysis in Practice*. 2nd ed. Boca Raton, FL: Chapman & Hall.

Guetzkow, J., M. Lamont, and G. Mallard. 2004. "What is Originality in the Humanities and the Social Sciences?" *American Sociological Review* 69 (2): 190–212. https://doi.org/10.1177/000312240406900203.

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

Haseman, B. 2006. "A Manifesto for Performative Research." *Media International Australia* 118 (1): 101. https://doi.org/10.1177/1329878X0611800113.

Hellqvist, B. 2010. "Referencing in the Humanities and its Implications for Citation Analysis." *Journal of the American Society for Information Science and Technology* 61 (2): 310–8.

Hellström, T. 2010. "Evaluation of Artistic Research." *Research Evaluation* 19 (5): 306–16. https://doi.org/10.3152/095820210X12809191250807.

Hicks, D. 2012. "Performance-based University Research Funding Systems." *Research Policy* 41 (2): 251–61. https://doi.org/10.1016/j.respol.2011.09.007.

Hug, S. E., M. Ochsner, and H.-D. Daniel. 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 (5): 369–83. https://doi.org/10.1093/reseval/rvt008.

Kälvemark, T. 2011. "University Politics and Practice-based Research." In *The Routledge Companion to Research in the Arts*, edited by M. Biggs and H. Karlsson, 1st ed., 3–23. London: Routledge.

La Porte, A. M., P. Speirs, and B. Young. 2008. "Art Curriculum Influences: A National Survey." *Studies in Art Education* 49 (4): 358–70. https://doi.org/10.1080/00393541.2008.11518747.

LERU. 2012. "Research Universities and Research Assessment." *League of European Research Universities*.

Lesage, D. 2017. "Against the Supplement. Some Reflections on Artistic Research." *FORUM+*. https://www.forum-online.be/en/issues/forum-maart17/tegen-het-supplement.

Lewandowska, K., and E. Kulczycki. 2022. "Academic Research Evaluation in Artistic Disciplines: The Case of Poland." *Assessment and Evaluation in Higher Education* 47 (2): 284–96. https://doi.org/10.1080/02602938.2021.1893651.

Lewandowska, K., Kulczycki, E., and Ochsner, M. 2023. "Evaluation of the arts in performance-based research funding systems: An international perspective." *Research Evaluation* 32 (1): 19–31.

McKee, A. 2020. "The Criteria Used by Key Decision Makers in Australia to Judge the Academic Quality of NTROs." *Media International Australia* 177 (1): 165–75. https://doi.org/10.1177/1329878X20921565.

Ochsner, M. 2022. "Identifying Research Quality in the Social Sciences." In *Handbook on Research Assessment in the Social Sciences*, edited by T. C. E. Engels and E. Kulczycki, 48–66. Cheltenham: Edward Elgar. https://doi.org/10.4337/9781800372559.00010.

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

Reeves, C. 2016. "Composition, Research and Pseudo-Science: A Response to John Croft." *Tempo* 70 (275): 50–9. https://doi.org/10.1017/S0040298215000625.

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 (4): 430–43. https://doi.org/10.1177/1474022218757151.

Rust, C., J. Mottram, and J. Till. 2007. *Review of Practice-led Research in Art, Design & Architecture*. Arts and Humanities Research Council.

Stankiewicz, M. A. 2000. "Discipline and the Future of Art Education." *Studies in Art Education* 41 (4): 301–13. https://doi.org/10.2307/1320675.

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

Trowler, P. 2013. "Can Approaches to Research in Art and Design be Beneficially Adapted for Research Into Higher Education?" *Higher Education Research & Development* 32 (1): 56–69. https://doi.org/10.1080/07294360.2012.750276.

Vear, C. 2022. *The Routledge International Handbook of Practice-Based Research*. London: Routledge.

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

Ysebaert, W., and B. Martens. 2018. "The ECOO M-VUB Stakeholder-driven Evaluation Design for Art and Design Research Outcomes." In *Evaluating Art and Design Research. Reflections, Evaluation Practices and Research Presentations*, edited by B. Van Kerckhoven and W. Ysebaert, 1st ed., 93–102. Brussels: VUBPress.