



Evaluation of Participants' Experience of the Quality Metrics National Test Phase

6th September 2016

by:

Nordicity



Table of Contents

1.	Executive Summary	1
2.	Introduction and Background	4
2.1	Arts Council England and the Quality Metrics Framework	4
2.2	Culture Counts Platform	5
2.3	Scope of the Nordicity Evaluation of the National Test Phase	5
3.	Methodology and Consultee Profile	6
3.1	Methodology and Limitations	6
3.2	Survey of Participating Organisations Profile	8
3.3	Survey of Peer Assessors Profile	10
4.	Key Findings	12
4.1	Effectiveness of the metrics	14
4.2	Suitability and viability of the platform	34
4.3	Capacity, usefulness and alignment	40
4.4	Overall attitude and propensity for future adoption	50
5.	Closing remarks	63
	Appendix A: Participating Organisations' Survey Report	64
	Appendix B: Peer Assessors' Survey Report	84
	Appendix C: Consultation List	92
	Appendix D: Bibliography	96
	Appendix E: List of Figures	98



1. Executive Summary

Arts Council England has, as part of its continuing investment in and support of the arts and culture sector in England, been considering the challenge arts and cultural organisations face in measuring and evidencing the quality of their work. This has been in line with its commitment to *excellence* as part of Goal 1 of its mission.

To achieve this multidisciplinary assessment of quality, Arts Council England has supported the sector to develop and test an open source set of metrics (Quality Metrics) via a proprietary digital platform (Culture Counts). The work takes its inspiration from a project initiated in 2010 by the Department of Culture and the Arts in Western Australia and has been developed over two pilot programmes in England prior to the current national trial, which has had 150 Arts Council England portfolio organisations testing the framework (of metrics and platform).

The Trial

The national phase of the trial was to be assessed by the platform providers and Arts Council England (the Arts Council); in addition, the Arts Council sought an independent third party evaluation of the participating National Portfolio Organisations (NPOs) and Major Partner Museums (MPMs) experience of the national phase. Nordicity was retained to evaluate the experience of participating organisations, identifying what had worked and not worked with the framework, and to assess the appetite for future adoption, including identifying challenges and opportunities.

The framework is constituted as a triangulated process, combining:

1. self-assessment (organisations setting their own parameters and assessing to what extent they have been met);
2. peer assessment (external assessment by peers applying their professional expertise); and
3. audience assessment (external assessment by the public).

To conduct the assessments, a set of qualitative statements are completed via a survey, the answers are ascribed a numerical value, which then provide a measure of the quality of the art or cultural activity being assessed. These data can then be accessed, compared and manipulated through the Culture Counts dashboard.

The Evaluation Process

To assess the participant's experience of the national phase, the research consultation comprised a survey of both participating organisations and peer assessors, in addition to a series of focus groups and interviews with stakeholders. These were supplemented with interviews of organisations that chose to not participate, organisations with a diversity focus, the developers of both the metrics and the digital platform, as well as potential alternative platform providers, and the Department of Culture and the Arts in Western Australia.



The Findings

The findings demonstrate that despite there being some misgivings conceptually about metricising an attribute such as quality, the sector sees value in the idea of being able to consistently measure quality across different events and across disciplines by any cultural or arts organisation, and is willing to work with Arts Council England to overcome shortcomings.

Simply having a national phase trial has resulted in a number of positive intended and unintended outcomes. Most importantly, it has initiated debate within portfolio organisations about defining and measuring quality, and about the use of data to support this. These discussions have occurred across departments and with both operational and strategic staff. Serendipitously, there have also been specific occasions where the trial has led to additional CPD opportunities for staff and to a deeper engagement with audiences (where both organisation and audience have benefited). Peer assessors nominated by participating organisations too have benefited, having been stimulated by experiencing art or culture from beyond their own related organisations or networks when conducting peer assessments.

The individual Quality Metrics (QM) themselves, which we refer to as 'dimensions' of quality, have previously been developed by the sector through two pilots and provide a basis for standardising quality measurement. The surveys indicated that both participating organisations and peer assessors were broadly positive about the individual QM dimensions themselves. However, there are some concerns about accessibility of language for non-specialists and a potential need for further refining of language with specific QM dimensions to better clarify what is meant and/or provide additional nuance or differentiation.

The idea of the triangulated evaluation process is seen as an important and innovative feature of the QM framework. However, operationally, it does present an appreciable administrative challenge to organisations to administer, though the survey element has been praised for being fast and easy to use. Importantly, there is evidence of misunderstanding by some organisations of the self- and peer assessment elements of the framework (in regard to a supposed bias and a perceived lack of flexibility) that unnecessarily impacted negatively on their perception of the QM framework.

In addition, there are specific challenges in regard to the process that are consequent upon the type of organisation or activity. For example, a touring operation experiences different challenges to a venue-based organisation, as does a one-off ticketed event compared to an un-ticketed exhibition (such as at a museum or gallery) lasting up to three months. Consequently, flexibility (or a recognition of limitations) is needed for any quality evaluative tool to be relevant to all organisations or disciplines.

There are some specific design concerns of the survey platform that were felt to introduce potential positive biases and some specific amendments were identified that would ensure wider accessibility but these issues could technically be readily addressed. However, despite efforts to the contrary, all surveys have some degree of bias. The need for additional kit (and cost, particularly for small organisations) and solid connectivity are more problematic. However, the current platform providers were widely praised for their responsiveness and effectiveness in addressing participants' process and platform problems.



Despite the desire in the sector for effective evidence-driven decision making, the consultation revealed two major capacity issues that undermined the use of data and a positive perception of the framework.

- i. Administrative challenge – the demand on organisations’ time/staff to properly administrate the QM framework was significant and in some cases had a detrimental impact on other activities (for example, it was commonly reporting that Audience Finder had to be deprioritised).
- ii. Skills gap – the familiarity and ability to effectively understand and exploit data was a challenge for the majority of organisations. In part, this led to some misunderstandings of the overall process, and sometimes it led to unduly resource-intensive analysis of the data, and/or an insufficiently rich understanding of the data produced.

Consequently, there is a need to increase the skills of practitioners within the sector and support the sector to find practical solutions to administrative challenges if a true shift in data culture is to be achieved. The former, of itself could alleviate some of the administrative challenge through more efficient operation of the framework and interpretation of the data.

It became evident that greater clarity and stronger messaging is needed about what is intended by the QM programme. Beyond the data competence challenges for organisations, there is an encompassing confusion about, and fear of, the ultimate purpose. This response is further calibrated by whether QM leads in participating organisations came from a marketing or programming background, as buy-in appeared to differ.

Organisations see that QM has the potential to become a useful evaluative tool for them to understand their audiences’ responses and the effectiveness of what NPOs/MPMs are trying to achieve. However, given the high-level, abstracted nature of the data and the lack of contextualisation, the majority of organisations do not see that it will be effective as a tool for benchmarking with other organisations and would fear it being used as such by the Council England or others. It is worth noting that participating organisations at the time of this consultation had not seen the aggregated dataset and analysis reported by CC.

There was also concern about data ownership and privacy; these concerns would dictate to what extent organisations would be prepared to share data, especially given the potential commercial sensitivity. Arts Council England would need to establish whether the data would fall within the ambit of Freedom of Information legislation.

The Future

To conclude, there is a desire for better data collection and use, and an interest in the possibilities presented by QM, but enthusiasm is measured and contingent upon a number of key amendments and assurances. To truly achieve a change in data culture, clearer articulation of the purpose of the QM programme is needed to ensure buy-in, and an upskilling of practitioners so that there is greater data competence, especially in regard to interpretation and analysis. This will take time, but the sector has shown willing and provided direction on how to get there.



2. Introduction and Background

2.1 Arts Council England and the Quality Metrics Framework

Quality Metrics (QM) is a sector-led and funder-supported evaluation framework that is intended to capture the quality of arts and cultural work by means of self, peer and public assessments. The framework takes its inspiration from a project initiated in 2010 by the Department of Culture and the Arts in Western Australia,¹ which commissioned consultants to work with arts organisations to develop a system that would help them understand the public value of arts and cultural activities.

It is of strategic importance to Arts Council England that the organisations it funds have a greater understanding of what their peers and audiences value about their work. It is also in the interest of these organisations to be able to use this understanding to uphold the quality of their work, to communicate that quality with their stakeholders in a more meaningful manner, and increase their overall resilience.

With this as a rationale, in 2013 Arts Council England supported Cornerhouse's request to trial Quality Metrics. This became the first QM pilot and involved a consortium of Manchester arts and cultural organisations testing the viability of the framework, developing a set of metrics and testing them across eight events.² The second pilot, funded through the Big Data strand of the Digital R&D Fund for the Arts and also led by the Manchester consortium, further developed the metrics and the digital platform.³

Following the two pilots that tested and developed the metrics with eight and 20 participating organisations, respectively, the Arts Council decided to offer a larger and more diverse number of portfolio organisations across England the chance to trial the framework.

Organisations wishing to take part in the trial submitted their interest in autumn 2015. A total of 150 organisations were originally accepted onto the national test phase. The national test phase ran from October 2015 to June 2016, with participating organisations evaluating programme activity between November 2015 and May 2016. A final report on the operation of the national test phase and the provision of an aggregated data set from the Culture Counts providers is expected by the end of June 2016.

¹ Public Value <http://www.dca.wa.gov.au/research-hub/public-value/>

² <http://www.artscouncil.org.uk/quality-metrics/quality-metrics>

³ <http://artsdigitalrnd.org.uk/projects/cornerhouse-et-al/>



2.2 Culture Counts Platform

The QM themselves are a core set of statements developed by arts and cultural organisations. These statements, which describe the components of quality artistic and cultural work, are referred to as Quality Metrics 'dimensions' in this report. The dimensions are open source and, for the duration of the national test phase, were administered on a cloud-based software platform called Culture Counts (CC),⁴ which provides organisations with a standardised means to collect and analyse self, peer and audience feedback on events, exhibitions or performances, and to share them across organisations. The CC platform and the technology behind it were provided by Counting What Counts (CWC).⁵

2.3 Scope of the Nordicity Evaluation of the National Test Phase

Nordicity was commissioned by Arts Council England to evaluate the experience of participating organisations and peer assessors over the course of the national test phase, including the CC platform. In addition, Nordicity's evaluation also included the views of a sample of portfolio organisations (NPOs/MPMs) that had chosen not to take part or had dropped out of the national test phase.

This evaluation does not cover the concurrent trial, where an additional 24 organisations are further developing and testing a set of metrics suitable for Children and Young People (CYP) and for Participatory work.

The Nordicity consultation was conducted in the latter stages of the national phase trial; consequently it is an evaluation of participating organisations' (NPOs/MPMs) and peer assessors' perceptions whilst the trial was ongoing and before the participants had had a chance to view or explore the aggregate dataset that resulted from the trial. Neither does this evaluation, because of time constraints, cross-reference its results with the aggregate dataset produced by CWC or the final report of the national test phase.

⁴ <https://culturecounts.cc/>

⁵ Counting What Counts Ltd. (CWC) is the limited company with the licence to operate the CC platform in the UK and Europe.

3. Methodology and Consultee Profile

3.1 Methodology and Limitations

The methodology for undertaking this evaluation comprised both primary and secondary research in the form of **desk research** and a far-reaching consultation through **two online surveys to over 300 stakeholders** comprising **over 200 stakeholder responses, discussion with 39 individuals** comprising **19 one-on-one interviews, four focus groups across the country with 20 individuals** and a **number of bespoke email questionnaires** followed by a process of **data synthesis and analysis**.⁶

Desk-based research

The desk-based research was conducted throughout the project and covering over **20 documents from 14 sources**; it was instrumental in informing both the design of the primary research methods and in synthesising and analysing the primary data from the surveys, interviews, focus groups and email questionnaires.

Consultation

The extensive consultation of **249 stakeholders** comprised (i) interviews, (ii) surveys, (iii) focus groups, and (iv) questionnaires, spanning the breadth and diversity of England's arts and cultural sector, including representation from:

- The nine regions across the Arts Council's five areas;
- Across the cultural disciplines and art forms;
- Organisations of differing sizes and with differing levels of investment from the Arts Council;
- Across the breadth of diversity in terms of age, ability/disability, ethnicity and religion, gender, sexual orientation;
- Both participating and non-participating organisations (incl. NPOs and MPMs);
- Both Organisation-Nominated (ON) peer assessors and AQA peer assessors;
- Diversity- and inclusion-focused stakeholders or those with a focus on the protected characteristics;⁷
- Metrics platform providers; and,
- Both domestic and international good practice stakeholders.

The consultation adopted a staggered and iterative approach, using early learnings to inform subsequent stages.

⁶ See Appendix C and D for the list of consultees and full bibliography, respectively.

⁷ 12 of the 19 participating organisations with a diversity or inclusion focus, responded to the consultation.

(i) Interviews

Interviews were conducted with **19 stakeholders** from across the breadth of the sector to focus on specific themes. Interviews were conducted with participating and non-participating organisations, peers and data suppliers. Themes addressed in the interviews covered motivations, audiences, diversity and accessibility, both the QM framework and dimensions themselves, data suppliers, non-participation, attitudes including propensity for future adoption, international experiences and strategic opportunities.

(ii) Surveys

Two online surveys were designed and distributed concurrently to a **total population ('N') of 316** contacts, receiving a **total sample size ('n') of 209** responses and accounting for an overall 66% response rate.

- Survey of Participating Organisations: One survey was distributed directly to a **population ('N') of the 144** NPOs/MPMs participating in the national test phase (i.e., the participating organisations), receiving a **sample size ('n') of 97** responses and accounting for a 67% response rate.
- Survey of Peer Assessors: The other survey was distributed directly to a **population ('N') of the 172** peer assessors participating in the national test phase, receiving a **sample size ('n') of 112** responses and accounting for a 65% response rate.

The surveys were open for a period of a month, from 22 March to 22 April 2016, with six reminders each sent in this time. Respondents across both surveys represented all of the cultural disciplines and art forms, English regions, size of organisation, level of investment by the Arts Council, and degrees of experience with the QM framework.

(iii) Focus Groups

A series of four focus groups were conducted with a total of **20 participating organisation and peer assessors** from across the cultural disciplines and art forms, English regions and degrees of experience with the QM framework.

The focus groups were carefully constructed to ensure NPOs/MPMs were represented from all disciplines, were of differing organisation size, had a breadth of experience of the process (i.e. high, medium and low engagement with the national test phase), and had differing levels of investment. They were hosted in accessible venues in Birmingham, London and Manchester. Template analysis was used to code the resulting qualitative data into hierarchical themes and sub themes.

(iv) Email Questionnaires

Throughout the evaluation, a number of bespoke email questionnaires were tailored to specific individuals and organisations to collect information focused on key issues as they arose in addition to other forms of consultation.

3.2 Survey of Participating Organisations Profile

Of the total population of 150 NPOs/MPMs participating in the national test phase, 97 completed the survey, producing a sample population of nearly two-thirds. The sample population is profiled below and it can be seen that a wide breadth of organisations contributed to the participating organisations survey: ranging across disciplines, geography, organisation size, the Arts Council funding levels and levels of engagement with the national test phase. In summary:

- All culture disciplines were represented (**Figure 1**).
- Nearly one-third were Theatre (32%) and nearly one-quarter were Combined Arts (24%).
- Literature and visual arts are least well represented as a reflection of the balance within the national portfolio, 2% and 8% of survey respondents, respectively, as opposed to 7% and 18% comprising the Arts Council's portfolio. Whereas, theatre and museums are over-represented at 32% and 11% of survey respondents, respectively, as opposed to 24% and 3% comprising the Arts Council's portfolio.
- In terms of size, as represented by the number of permanent staff, responses were roughly evenly divided⁸ among small, medium and large organisations.
- The location of respondent organisations was across England: ranging from the North (35%) and London (28%), to the Midlands (8%), the South East (16%) and South West (12%).
- Three-fifths of respondent organisations had held three or more QM-assessed events (60%) with nearly a further two-fifths having conducted one or two (38%). ()
- Respondent organisations represented NPOs/MPMs receiving differing levels of investment from the Arts Council –45% reported they had received under £250 thousand and a further 40% reported having received £250 thousand to £1 million. The remaining 15% reported more than £1 million in investment.
- Over one-third of respondents had also acted as individual peer assessors for another arts or cultural organisation, and so had gained direct insight into the application of the QM dimensions to activities that were being assessed as part of the national test phase.
- There were 17 diversity-led NPOs/MPMs (or those with a diversity focus within a wider remit) participating in the national test phase, of which 11 responded to the survey.
- The primary motivation for participating in the national test phase by nearly half the respondent organisations was the desire to improve their own quality assessment. In addition, nearly a quarter were specifically interested in the QM concept. The Arts Council

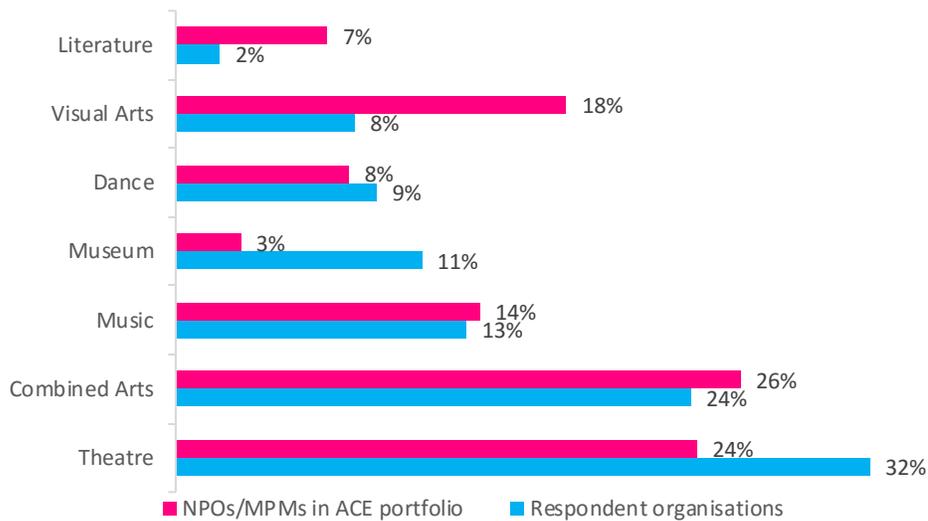
⁸ ±5 percentage points.

encouragement was the second strongest primary motivation for over a third of respondent organisations.

- Data sharing was not a primary motivation for any of the respondent organisations, and for a large majority (65%) of the sample population it was the least important of the suggested reasons for participation provided for respondents to consider.

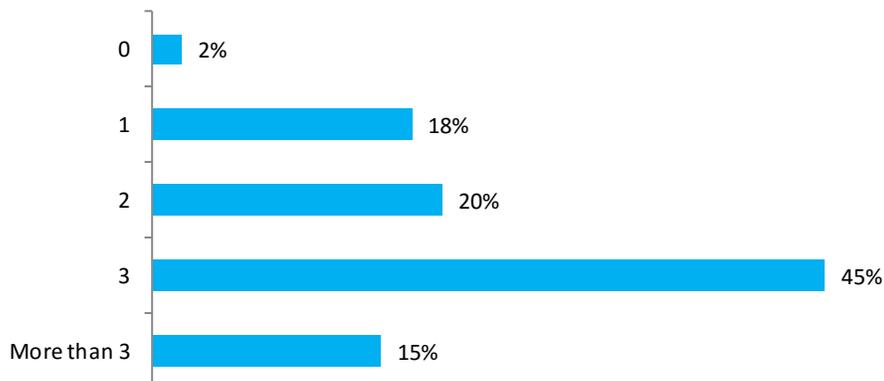
The full Nordicity Survey Report is included in the Appendices.

Figure 1: Respondent organisations' primary discipline



Count = 688, 97

Figure 2: Number of events assessed by respondent organisations for national test phase



Count = 97

3.3 Survey of Peer Assessors Profile

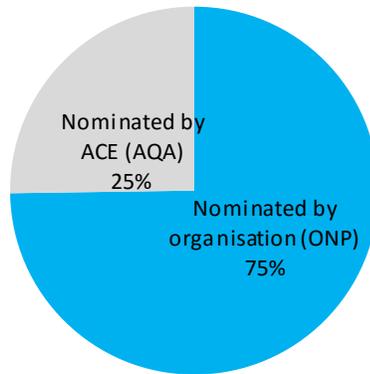
Of the known total population of 172 participating peer assessors in the national test phase, 112 completed the survey, producing an estimated nominal sample population of nearly two-thirds. However, it was clear from the focus groups and survey responses that participating organisations recruited assessors that were in addition to, or instead of, peers on the list provided, and so it is not possible to say what the true total population size is, and therefore, the true response rate.

The sample of peer assessor respondents is profiled below:

- Of survey respondents, 75% had been organisation-nominated and 25% had been the Arts Council-nominated. The latter were Artistic and Quality Assessment (AQA) assessors. **(Figure 3)**
- Responses were received from peer assessors from all disciplines but those from Theatre, Combined Arts and Visual Arts numerically dominated all the other disciplines.
- The high frequency of combined arts peer assessors (26%), would suggest a broader base of expertise across more disciplines, as 'combined arts' are multidisciplinary.
- There was a slightly broader mix of disciplines amongst the ON peer assessors than the AQA peer assessors, the latter of which were dominated by nearly two-thirds coming from Theatre or Combined Arts. **(Figure 4)**
- There was a broad mix of experience of using QM framework and the CC platform, ranging from 11% of respondents having assessed more than three QM-assessed events to nearly one-fifth (18%) who had not yet assessed any. The largest cohort of 37%, had assessed one event at the time of the survey.
- Events assessed were dominated by the disciplines of Theatre, Combined Arts and Visual Arts, which together constituted 75% of all events attended by survey respondents.
- Location of events assessed by survey respondents was spread across the five Arts Council areas, though the North and London dominated.

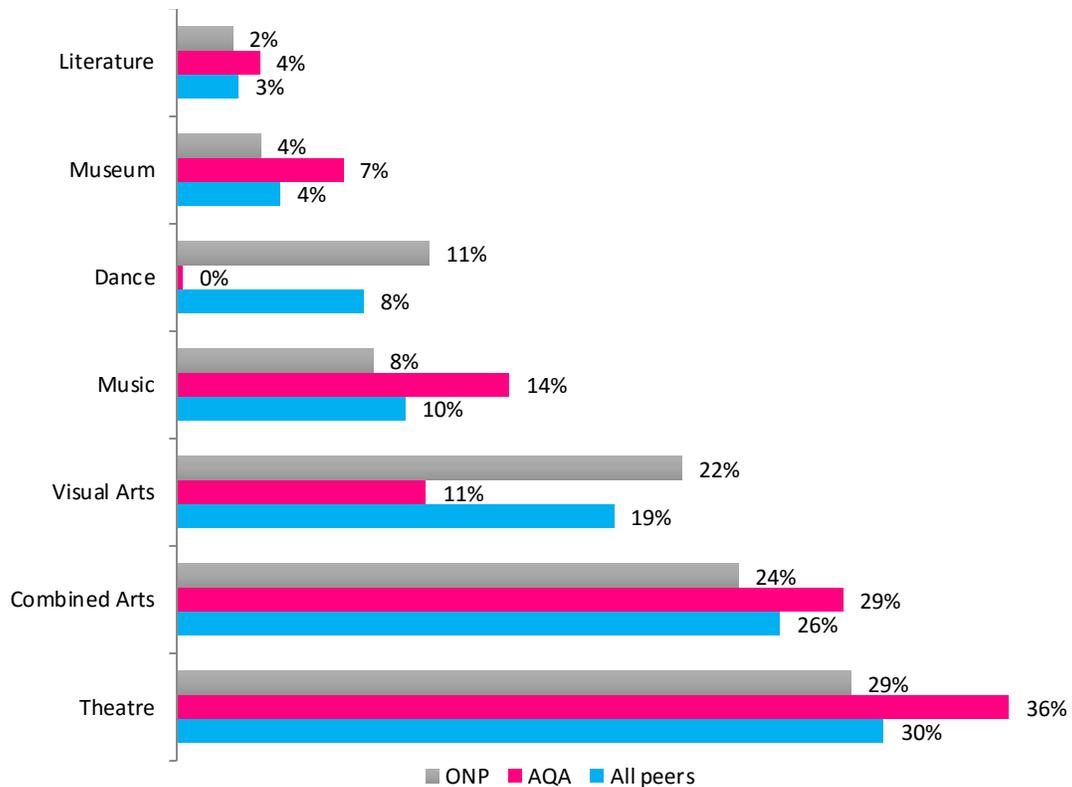
The full Nordicity Survey Report is included in the Appendices.

Figure 3: Type of peer assessor



Count = 111

Figure 4: Primary cultural discipline or area of expertise of peer assessors



Count = 83, 28, 112

4. Key Findings

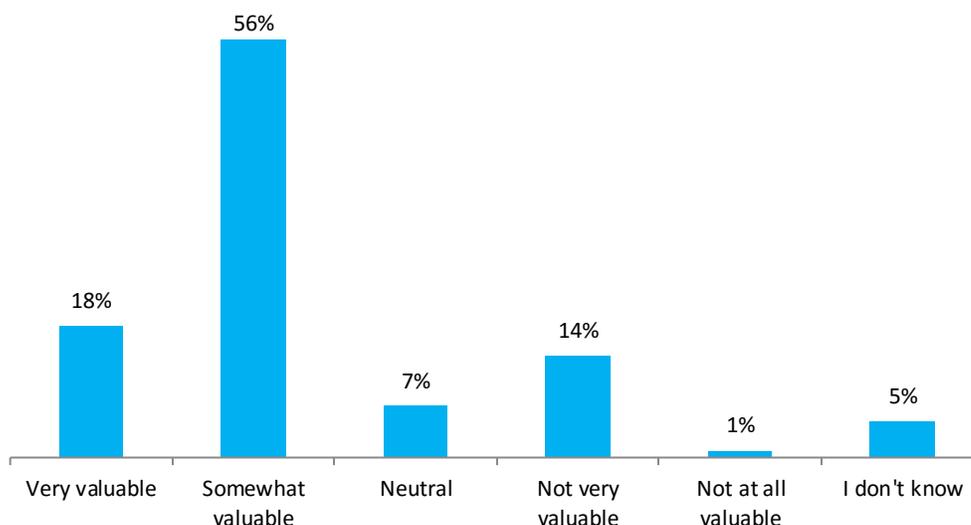
The commentary below presents insights from Nordicity’s evaluation of the national test phase of the QM framework. It is a considered analysis across all three strands of the investigation (survey, focus group and interview), unless stated otherwise. Telling the story from the perspective of organisations and peers, this section is organised in four parts that examine (i) the effectiveness of the metrics, (ii) suitability and viability of the platform, (iii) capacity within the NPO/MPM sector and potential for alignment with their existing systems and processes, and, (iv) appetite for future adoption.

The consultation resulted in a widely answered survey, achieving a two-thirds response rate from the participating NPOs/MPMs, and which demonstrated the sector’s inherent interest in the QM framework overall and in its advancement.

Overall, survey respondents were broadly positive about the QM framework. However, whilst the higher level survey questions did not directly reveal specific challenges with the QM framework, these were explicitly raised in the survey open text data, focus groups and interviews. The quantitative and qualitative findings therefore present a complementary picture.

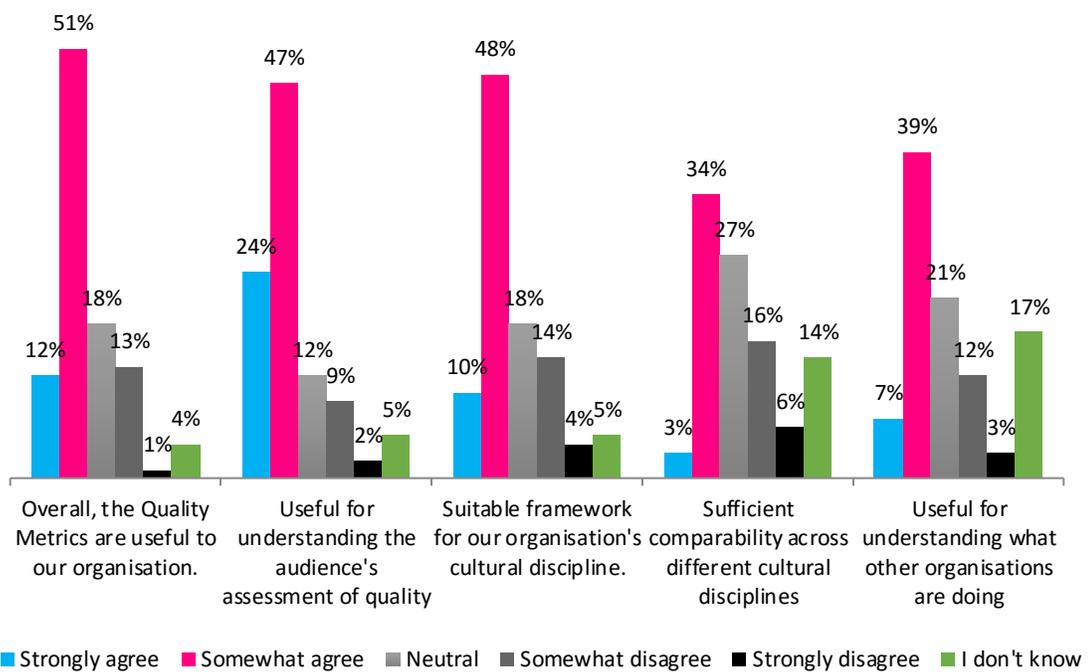
It should also be noted that as one of the primary research objectives of the consultation was to help the Arts Council understand how the framework could more readily meet the needs of organisations, there has inevitably been a greater focus on challenges and improvements. However, NPOs/MPMs did see value in the QM framework, viewing the national test phase as a positive step which could lead to useful change, as the following headline survey results indicate. (Figure 5 and Figure 6)

Figure 5: Value of QM concept to respondent organisations



Count = 102

Figure 6: What respondent organisation thought of particular elements of the QM framework



Count = 98

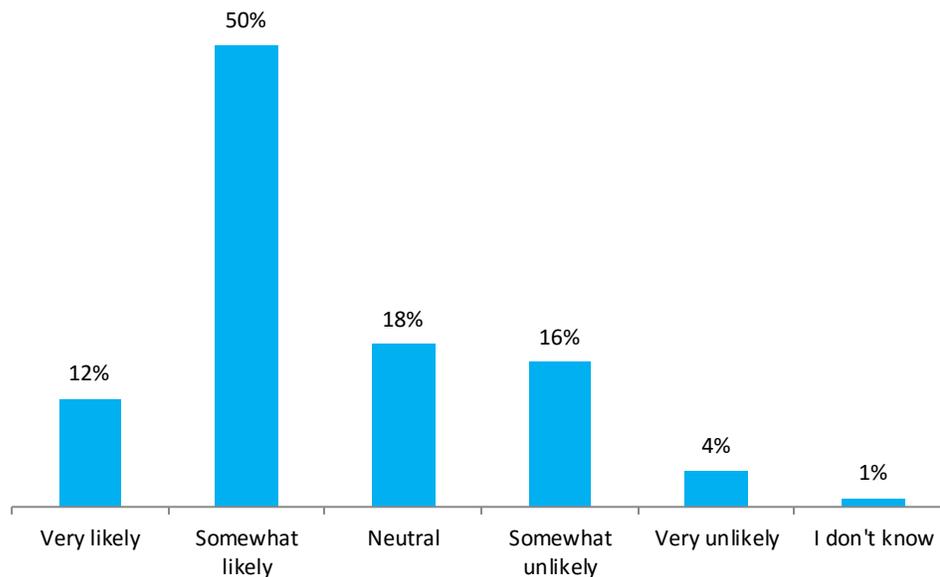
4.1 Effectiveness of the metrics

This section reports on whether consultees believed that the QM dimensions and overarching framework is an effective way to assess artistic quality, for their own organisation and for peer NPOs/MPMs. It is split into two areas: the first explores the individual dimensions specifically, and the second focuses on the overall evaluation framework.

i. The Quality Metrics dimensions

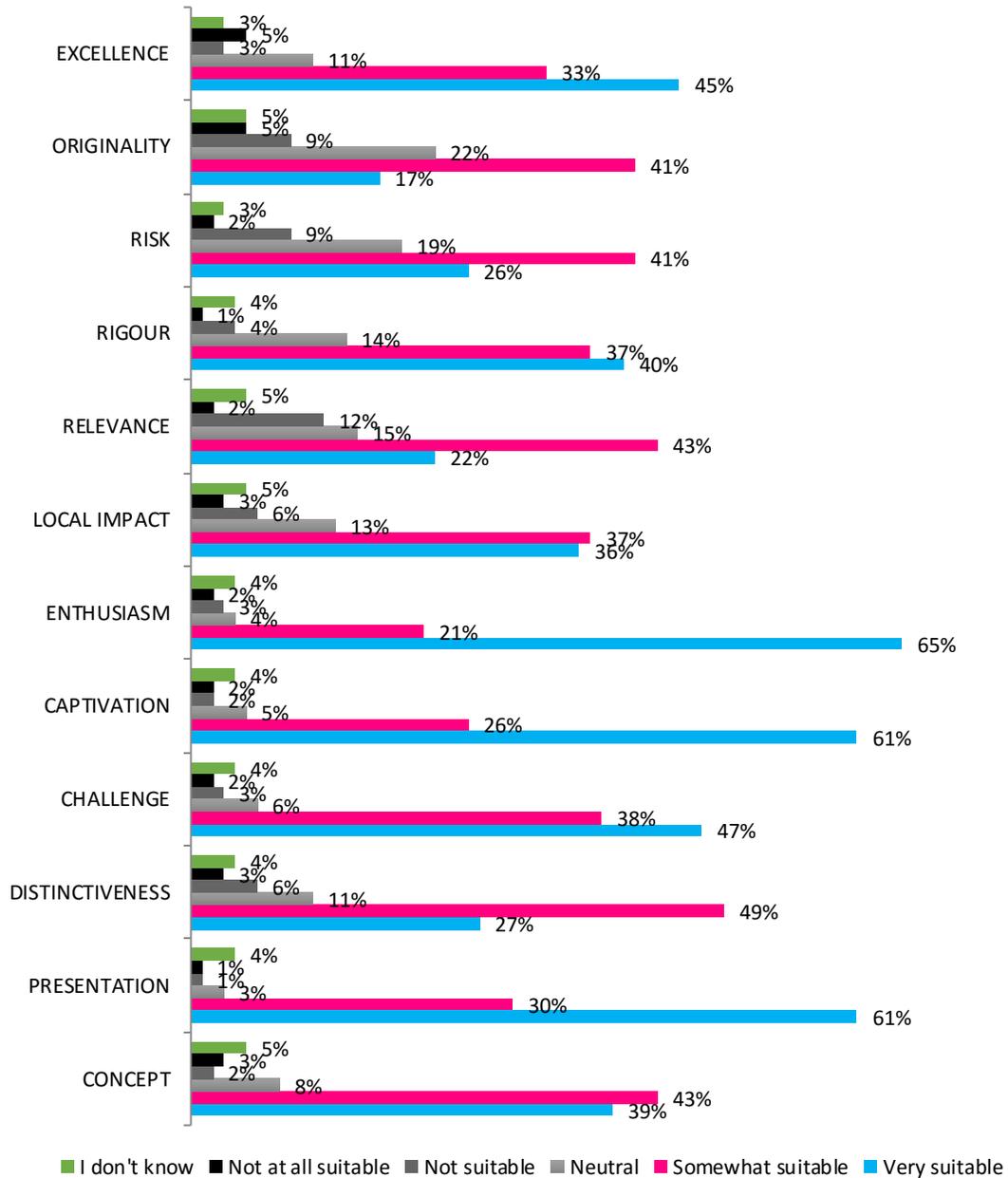
Overall, survey respondents were broadly positive about the various QM dimensions. Approximately one in eight respondents thought it very likely that their organisations would use these dimensions in the future, with a further half of respondents thinking it somewhat likely, resulting in a total of 62% believing that their organisation was very likely or somewhat likely to use the QM dimensions in the future (Figure 7). Three dimensions in particular were identified as being very suitable for measuring quality by the majority of respondents: *Enthusiasm*, *Presentation* and *Captivation* (Figure 8).

Figure 7: Could respondent organisations anticipate their organisations using the QM dimensions in future



Count = 102

Figure 8: Suitability of individual QM dimensions for measuring quality

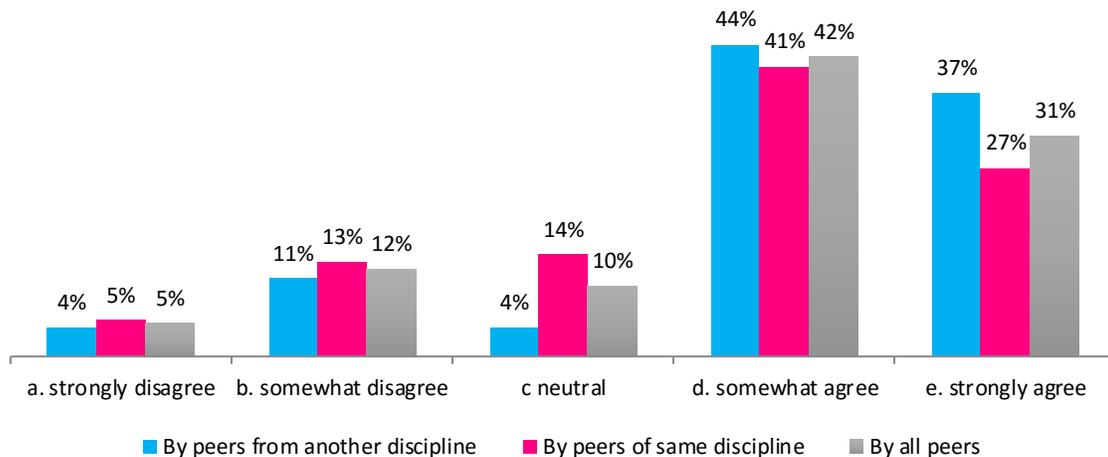


Count = 98

It is important to note here that among the 11 NPOs/MPMs comprising the diversity cohort⁹, there was a significantly stronger likelihood of using the dimensions – nearly one in five thought it very likely and a further near three quarters thought it somewhat likely, i.e. 91% thought their organisation was very or somewhat likely to use the QM dimensions in future. (Figure 45)

Peers were also broadly positive towards the dimensions but less so than the participating organisations. When considering the specific events that peers had assessed, over three in 10 strongly agreed that the dimensions were suitable for assessing the event, with a further four in 10 somewhat agreeing. When this is split by peers of the same discipline or a different discipline, peers of a different discipline agreed more strongly and those of the same discipline less strongly. (Figure 52) Note that over 40% of these events were from a single discipline, namely, theatre.

Figure 9: Suitability of QM dimensions for assessing the event attended as judged by type of peer (of the same or different discipline)



Count = 73, 98, 171

There is also qualitative evidence to indicate that certain dimensions are valued by organisations, for example some organisations have been motivated to incorporate the metrics into their internal evaluation or monitoring systems.

Quote 1: “I found out quite recently that one of our project teams had actually adopted three or four of the metrics for their own feedback that they collect so there is value there...” (Organisation)

⁹ NPOs/MPMs either diversity- or inclusion-led or with within a wider remit participating in the national test phase.

Prior to the national test phase, the dimensions were developed and tested by a consortium of eight Manchester arts and cultural organisations. Further refinement and development of the metrics followed, through the Big Data strand of the Digital R&D fund for the arts, also led by the Manchester consortium and tested with 20 organisations¹⁰. Despite this, whilst the quantitative data suggests a general positivism overall towards the dimensions, deeper scrutiny revealed there were a number of common refinement suggestions made by both peers and organisations through the qualitative data to make the metrics more effective. These focused on the following five core areas of difficulty.

1. Not all metrics were viewed as appropriate measures of quality

This perception was most commonly reported with *Risk*, *Local Impact*, *Challenge* and *Distinctiveness*, which both organisations and peers suggested were more likely to be overall Arts Council objectives or descriptors rather than true indicators of quality. One of the main recurring suggestions of 'missing' dimensions was a statement to capture quality in relation to accessibility and diversity. Consultees also queried why audiences were not required to answer the three additional questions for peers and self, stating that these questions are equally as relevant to the public (particularly *Excellence*), mainly given that the incremental completion time would not be significant. It should be noted that if organisations wanted to include these additional dimensions, the software does allow this.

Quote 2: "...some of the questions feel to me that they're answering Arts Council objectives. I think Distinctiveness might be one. Local impact is crucial for Arts Council but for punters, I'm not sure whether they'd know what that would mean. I think there might be a hint of a tension somewhere." (Organisation)

Quote 3: "...The qualities measured seem to exclude as much as they reveal e.g., as well as 'Empathy', why not 'Fun', 'Stimulation', 'Emotion', 'Understanding', 'Encouragement', 'Empowerment', 'Heritage', 'Community', 'Inclusivity', etc...? The measures chosen actually seem quite institutional and conservative in their scope." (Organisation)

Quote 4: "I went to [music performance] and it was definitely different to things I've experienced before, but that's a completely different issue to whether I think it's any good or not. They're descriptors and often it's quite interesting but shouldn't be seen as demonstrating quality." (AQA Peer)

Quote 5: "The creative value of diversity is not reflected in the questions, even though this does have a significant bearing on the quality of artistic product (cf. the Arts Council's Creative Case for Diversity). It would be good to think about how this could be reflected in the Quality Metrics." (Organisation)

¹⁰ There is a wide range of literature publicly available which reports on, and/or evaluates these two previous pilots, and the original concept developed in Western Australia in 2010. The literature review list is provided in Appendix D.

2. Semantic variation potentially affects comparability of data

The semantics of the dimensions were called into question, particularly with *Local Impact (It's important that it's happening here)*, where 'here' could be interpreted in different ways (and has been by the various consultees taking part in this evaluation, thus impacting the comparability of resulting data).

Similarly, *Distinctiveness (It was different to things I've experienced before)* was perceived as problematic by both organisations and peers in focus groups. For example, it was reported that the dimension may be interpreted differently by a first time attender, as opposed to a regular arts attender. The semantics of metrics such as *Distinctiveness*, as well as *Originality* and *Captivation*¹¹ (*It was absorbing and held my attention*), were also deemed problematic and open to multiple interpretations. This finding suggests that further strict clarification and definition of meaning may be required to ensure that self, peer and public assessments are interpreting each metric consistently and objectively.

Quote 6: "It was interesting when we were talking to audiences about Captivation and they immediately said 'well yeah I got my money's worth.' That was their immediate response to those things. When actually that's not what that metric is driving towards." (Organisation)

Quote 7: "I have a problem with that [local impact] question because of the word 'here'. What does it mean? Here as in this venue, or here as in this borough, here as in London..." (Organisation)

Quote 8: "One person's ground-breaking is another person's average day at the office." (ON Peer)

Quote 9: "The tricky one is Distinctiveness, it's tricky if you're walking into a music gig what are you asking there? Is it the content? What do you want me to assess here?" (Organisation)

3. The language of the metrics was not always accessible to all

From an accessibility perspective, some organisations reported that they had to proactively support staff in understanding the terminology and meaning of each metric to enable them to undertake fieldwork with audiences. Indeed, "abstract" and "sector speak" were frequently used by consultees to describe the dimensions. In cases where the survey had been emailed to audiences there is some evidence to indicate that the public had difficulties with understanding the questions.¹² There is also an indication that those with learning disabilities, learning difficulties and audiences with limited English or literacy skills would need additional support to effectively complete the survey.¹³ These findings indicate that further interventions in terms of the methodology or process would be required to ensure that the survey does not exclude certain voices.

¹¹ Note the inherent variability in interpretation, as *Captivation* had been identified in the survey as one of the more suitable measures of quality (see beginning of 4.1 i).

¹² The bespoke challenges faced by particular types of organisation are discussed further in Section 4.1 ii.

¹³ Other accessibility challenges exist with the survey platform and are discussed further in Section 4.2 i.

Quote 10: “Access is an issue for speakers of other languages or people less used to a broad vocabulary – people struggled to understand some of the words.” (Organisation)

Quote 11: “In the ‘any other comments’ box at the end of my surveys I’ve had a high number of people commenting on the survey rather than the performance and going ‘I didn’t really see the relevance of the questions’ or ‘I didn’t get the questions’ – that’s the public being asked ‘do you have any other comments’ and saying ‘I don’t really see the relevance of the survey to the performance that I saw last night’” (Organisation)

4. Certain metrics are not sufficiently differentiated or lacked nuance

Some consultees reported that whilst *Rigour* and *Presentation* are two very different and relevant indicators of quality, their qualifying dimension statements (*It was well produced/presented* versus *It was well thought through and put together*) were too similar and therefore not distinctive enough to differentiate between the two dimensions. Therefore, some minor adjustments to individual dimension statements are likely to be required to ensure greater clarity and true differentiation. There is also a suggestion that peers (particularly AQA peers) required more technical nuance than self and public assessors; that the current dimensions are simply too vague and do not allow for ‘technical’ aspects of quality to be reported as fully as peer assessors wish. This suggests a seemingly contradictory view – that of simplifying the language (especially for public and self) and the need to create more nuance (especially for experienced peers).

Quote 12: “If you put those to an audience member you’d be hard pressed to get them to find the difference [between Rigour and Presentation].” (Organisation)

Quote 13: “When I was a peer I started feeling a bit frustrated with it. I found it quite hard to answer the questions as they were so vague and generic.” (ON Peer)

Quote 14: “... if you were to unpack excellence, you would be saying, there are two components if you want to crystallise it... would be, interpretation and execution, so a good performance will have both of those in high quality. You can have just one and not the other, so very fine insightful thoughtful interpretation but full of wrong notes, and vice versa, but there’s nothing there – excellence is rather generic and doesn’t explicitly unpack that – and I think it would be helpful if it did.” (AQA Peer)

5. One size does not fit all: organisations want greater flexibility

For the national test phase it was determined that to be most effective at trying to understand NPOs/MPMs’ experience of QM, a standardised set of core QM dimensions would be mandatory to facilitate a consistent test of the metrics at scale and gain insight into the cohort’s impression of them, in the full knowledge that the CC platform did allow for customisation and a degree of flexibility for participating organisations. In addition, the self-assessment does enable organisations to weight each of the metrics themselves prior to the event based on its unique aims and ambitions.

The purpose of making all of the QM dimensions mandatory in the trial was to stimulate debate and discussion around each one, based in part from the learnings of the experience in Western Australia. The rationale for such standardisation was not fully understood by some organisations, and the utility of the pre- self-assessment was not fully appreciated, both of which had a negative impact on their experience with the trial.

Despite mixed views about the purpose¹⁴ of the QM framework (i.e., whether it is predominantly an evaluative tool for critical reflection or wider benchmarking) the most commonly reported criticism of the existing dimensions was the lack of flexibility to tailor the evaluation to the needs of the participating organisations.

Whilst consultees predominantly understood the concept of the framework, the role of the self-assessment prior to the event as a means of weighting each dimension individually was not as clear. Some suggested it would be more useful to select indicators that reflected their own artistic objectives and intentions, rather than assess aspects of quality that are not intentional (e.g., assessing a pantomime against the *Originality [ground-breaking]* metric when it had no intention of being ground-breaking). Although there was a recognition that this would potentially reduce the ability to benchmark across the portfolio, it is evident that a 'measure what matters' approach may gain more buy-in for the framework from the sector.

It was regularly suggested that a small number of compulsory metrics could be chosen by the Arts Council, with a suite of remaining separate metrics available that organisations could select (potentially provided as templates per art form). Furthermore, consultees posed that Arts Council Relationship Managers could work with organisations when setting artistic objectives to ensure both Arts Council and the participating organisations were in agreement about the selection of metrics.

Quote 15: "I really like the idea of this but I want my own freedom to set my own benchmarks of excellence and quality and ask the questions that are truly going to resonate with my own audience." (Organisation)

Quote 16: "I'm not sure about measuring things that you're not trying to achieve." (Non-participating organisation)

Quote 17: "We'd like to be able to pick out the questions that are truly relevant to the event that we're putting on." (Organisation)

Overall, organisations commonly suggested that there is not a 'one size fits all' solution to the metrics due to the variety of disciplines and art forms in the portfolio, the individual differences in objectives as outlined, and the varying target audiences. Concern was raised over the relevance and comparability of the existing dimensions as a result.

Quote 18: "The metrics were very 'one size fits all' and not always suitable to the type of event programmed." (Organisation)

Specific examples were given by two different consultees (one participating and one non-participating) to demonstrate that the overall evaluation framework did not 'fit' their particular discipline. For example, a literature organisation delivering a product such as book gifting or reading support groups for parents through gatekeeper partners suggested that the existing wording of the QM dimensions was

¹⁴ Purpose is further discussed in Section 4.4.

not relevant. Similarly, a producing organisation touring work elsewhere in the country questioned the 'local impact' metric and usefulness to their own learning.

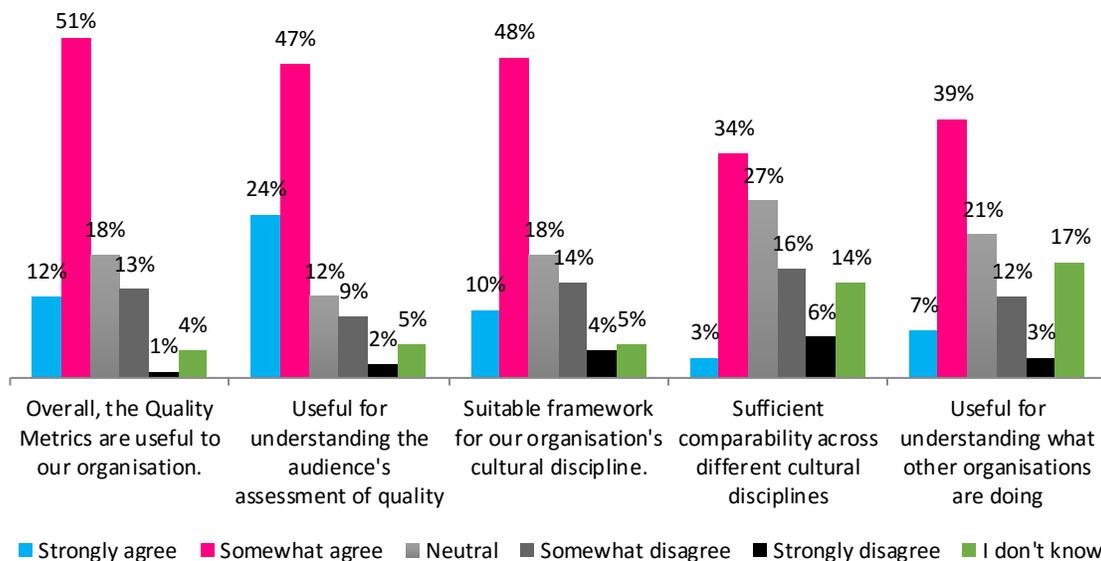
Quote 19: "I would come to something like this again' – that kind of wording is very events based isn't it?" (Non-participating organisation)

Quote 20: "The local impact question is problematic for our touring productions. That question provides some insight but it's one for the venue." (Organisation)

ii. The Quality Metrics framework

Overall, a majority (63%) of respondent organisations either strongly agreed or somewhat agreed that the QM framework was useful to their organisation (Figure 10).

Figure 10: What respondent organisations thought of particular elements of the QM framework



Count = 98

The survey results also indicated that respondent organisations thought that the framework would be useful for understanding their own audiences and that it was appropriate whatever the organisation's cultural discipline. In addition, there was also a positive, though less strong, view that the framework could be used to compare across disciplines and to understand other organisations. Despite this general positivism towards the framework at a headline level, the consultation did highlight a number of perceived challenges. The rest of this sub-section addresses these challenges.

Sampling and robustness requires review

In addition to the specific concerns around the QM dimensions, the majority of consultees questioned the reliability of the resulting data because of the sample frame, in terms of its representation and size.¹⁵ This aspect evidently impacted the organisations' use of the data, with organisations unconfident to draw any firm conclusions, unable to 'convince' programmers of its value, and unsure of what 'robust' would look like in practice.¹⁶ These issues relate to the terms of the trial rather than the framework itself. As the main purpose of the trial was to give a large number of organisations the opportunity to trial the framework, therefore to make this more manageable, organisations were encouraged to achieve a minimum of 30 public responses per evaluation, however, there were no restrictions upon them collecting more or in constructing representative samples of their audiences.

Furthermore, non-participating organisations cited the target sample for survey responses as a barrier for participation, particularly within the time period of the national test phase. Conversely, the target sample sizes were a benefit to participating in the national test phase for some organisations who found the public assessment quota of the QM framework much easier to achieve than that of Audience Finder. However, each tool is based on a different methodology, consequently each is subject to different participation constraints. The best practice research standards that were used to develop the Audience Finder methodology require the thresholds set, whereas, the QM methodology requires a relatively low minimum of responses for the term of this trial. However, participating organisations were given more detailed information about thresholds and sample sizes. This finding suggests that a careful balance will need to be found within the framework to provide datasets that are trusted, whilst simultaneously responding to the capacity challenges faced by organisations. It also suggests that further clarification of messaging around the use of the data and definitions of robustness may be required for those with less expertise in this field, as some of the quotes reveal.

Quote 21: "People that are self-selected have got something to say but it's not representative of the entire audience." (Organisation)

Quote 22: "I have no confidence at all in the self or peer assessment as data because you're asking so small a group. So I don't see how quantitatively asking five people is going to give any data, which is meaningful or robust. It might be interesting but it's not robust." (Organisation)

Quote 23: "I think there was very low morale across everybody working around Audience Finder because we just couldn't meet the targets. At least this felt achievable and morale increased." (Organisation)

Misunderstandings that need to be allayed, centre on the need for a robust sample size, which only applies to audience assessment (i.e. achieving a minimum number of returns), whereas, with the self-

¹⁵ For the national test phase, organisations were asked to survey three events, with a minimum 30 audience members assessing each and five peers.

¹⁶ Usefulness is further discussed in Section 4.3.

and peer assessment it is the difference in pre- and post-assessments or the quality of the assessor/assessment that is more pertinent.

Survey fatigue poses a risk and challenge for participating organisations

The findings indicate there are widespread concerns regarding the impact of survey fatigue – for public, self and peers. These are concerns common to all surveying, as the public's views and information are increasingly sought. It was perceived that continually repeating the survey might negatively impact upon: (i) the actual results (e.g., less care and thought would be taken during completion due to the 'boredom' factor), (ii) the reputation of the organisation (e.g., becoming known for 'bombarding' audiences with survey requests), and, (iii) staff time (e.g., having to create a new survey for each event – in comparison with the ongoing version of Audience Finder).

One solution posed by a number of consultees (particularly Combined Arts and Theatre organisations) was that to assessing the quality of a whole programme or season. This solution would address sampling difficulties, allow for a more meaningful or holistic view of activity overall, and avoid polarisation (e.g., when a cutting-edge contemporary art piece is assessed one week but not a commercial touring comedian who is performing the next).

Quote 24: "A few of our artistic people did the survey three times and they said they didn't think about it as much by the third one. The first time they did it, because it was so new, they were questioning it more deeply." (Organisation)

Quote 25: "...we have people who come 4 times a year, there's a lot of repeat bookers for a lot of different art forms. Soon they'll become very bored with us asking them these questions and on top of that we'll be asking them other questions. They'll soon get really cheesed off." (Organisation)

Quote 26: "The whole offer is a better assessment, because that's how we programme. We programme some commercial stuff, we programme challenge stuff, and overall that creates our programme. We hope people will engage across the range and that's where I feel it's sustainable, it's less vulnerable to be picked apart and we'd be more likely to engage with it." (Organisation)

Bespoke challenges exist for particular types of organisation

The experience of the national test phase of the QM framework presented a number of specific challenges for different disciplines and organisation types. Some of these challenges would be faced by other similarly constituted data collection exercises such as surveying. These included the following.

- ticketed organisations were unsure if they had gained a representative sample due to sending the survey via email to their bookers, whereas non-ticketed organisations using fieldworkers were able to select a random sample. Conversely, non-ticketed organisations experienced capacity challenges around conducting face-to-face interviews¹⁷ whereas ticketed organisations were able to use their box office database.

¹⁷ Capacity and administration burdens are discussed further in Section 4.3.

- ticketed organisations/organisations surveying at ‘one-off’ events faced more time pressures to collect survey data from audiences than non-ticketed museums and galleries surveying across longer periods (e.g., of an exhibition run). Similarly, organisations running smaller events found it difficult at times to achieve the minimum audience number required by the national test phase.¹⁸
- producing organisations touring their work to receiving venues could not access email data to manage the survey circulation and had no capacity to send their own fieldworkers.¹⁹
- ‘well-known brand’ organisations were concerned about higher expectations on them affecting assessment results (especially comparing pre- to post-assessment) versus challenges with ‘lesser known brands’ who struggled to attract peer assessors.²⁰

Quote 27: “With presenting partners – we’ve learned it needs to be instigated by them. So that’s taught us about the different ways to measure quality – trying to do it one step removed is too challenging. When you’re a step away from it, you don’t know how people are being asked, we’re totally out of control.” (Organisation)

Quote 28: “I think there’s a problem of our brand being so strong and well-known and that raises expectations ... but our artists take risks too...” (Organisation)

Quote 29: “Some of our events are quite small ... the target is often quite punishing if you’re only got 40 attendees and you’ve got to catch them when they come out. It needs to be so much more useful than it is at the moment for us to put all that investment in.” (Organisation)

The ability to add context is an essential requirement for organisations and peers

This was the most frequently referenced feedback across the whole consultation. Lack of ability to add any context to either the front-end survey to inform peers about the intentions and objectives of the piece, as well as to add to the back-end dashboard to support colleagues in understanding the resulting scores, it was seen as a missed opportunity. This perception was a misconception, as it is possible to write an introduction for public or peer assessor surveys, and so, this is another example of where improved communications are needed with the sector to avoid such misunderstandings. Being able to ‘tell the story’ was key not only for organisations to ‘justify’ scores but to provide a legacy for new staff coming into the workplace once those involved had left. It was also seen as vital to demonstrate clear reasoning behind selecting the work to be assessed, and to give context to the Arts Council.²¹

Both ON and AQA peers reported that they would have preferred more context about the organisation beforehand in order to efficiently and effectively score what they were assessing. Similarly, peer assessors (both ON and AQA) wanted to be able to provide their own narrative context alongside their

¹⁸ Sampling challenges are discussed earlier in Section 4.1 ii.

¹⁹ Indeed, this was also raised by non-participating organisations as a barrier to participation in the national test phase.

²⁰ Challenges with peer recruitment are covered later in this section.

²¹ Context is linked to a greater understanding of purpose and is discussed further in Section 4.4.

scores – in order to justify and explain their reasoning. Travelling to assess an event and then taking five minutes to complete a QM survey was deemed to be ‘short-changing’ the organisation being assessed.

Quote 30: “I would have liked a free style paragraph, that would have been really helpful, it’s fine to say this was doing important things, or this was revelatory, but actually why is it? There’s no opportunity to say what struck a chord and why ... and I think if people have committed to give an evening, people are willing to give a half decent report and give that detail.” (ON Peer)

Quote 31: “I think if I’d travelled any distance, even to make a special journey to see something and all I did was to do that form, I think I’d be short-changing them as much as me ... for five minutes! I don’t think it’s sustainable in a financial sense, you’ll be billing people for travel and that’s all they get out of it.” (AQA Peer)

Quote 32: “I think what might be missing is context – the context in which the piece was created was essential for your understanding of what you were seeing. You would have a different set of criteria and expectations if you went into that performance not knowing the context and circumstances in which it was created.” (AQA Peer)

Quote 33: “I think it’s difficult to answer some of these questions without reference to what the purpose or intention or aspirations of the organisation would be.” (ON Peer)

Bias or skew is potentially introduced by particular elements of the framework

Throughout the evaluation process, consultees suggested a number of areas where unintended bias or skewed data had the potential to be introduced. It is evident that these elements contributed to consultees’ overall opinion that the resulting data did not accurately reflect the quality of their work. There were five reoccurring identified areas of concern.

CC as part of their evaluation of the national phase have been considering the potential for bias and will be reporting separately on their investigations, which may shed light on the issues below.

1. Self-assessment

Consultees commonly reported that self-assessment can introduce bias through subjectivity, suggesting that it was difficult to remain impartial despite being briefed (or briefing colleagues) about the purpose and intention of the evaluation. Organisations often found that the self-assessment was scored overly highly, as assessors had prior knowledge/background about the work and the context. These observations are real but reveal a misunderstanding about the purpose and use of self-assessment. It is a tool for self-reflection by organisations, and as such, the differences between the pre- and post-assessments are what is crucial, not the strength of the assessment, thereby obviating any unintentional bias as a result of familiarity for example. Also, self-assessment is important as part of the understanding imparted by all three parts of the triangulation but the scores from each is not additive, as feared by organisations. As such, this is another example of the need for clearer and stronger communications about QM.

Quote 34: “It’s the self one I find...they’re always high, it’s the people who are producing it, putting it all together...so they expect it to be great, amazing and there’s high scores consistently.” (Organisation)

Quote 35: “... the self-assessor knows what’s gone into it and all the thought processes that the public doesn’t appreciate.” (Organisation)

2. Peer assessment

When organisations struggled to recruit peers,²² well-known external colleagues were often brought into help instead, which potentially introduced more favourable results.

Quote 36: “My peer assessors said ‘of course I’ll give you the right answers’ [wink].” (Organisation)

There is also some evidence to indicate that ON peers found it difficult to remain impartial when assessing organisations with whom they had previously partnered or worked with. In contrast, AQA peers have been ‘trained’ by the Arts Council through the AQA development process in an attempt to allow them to remain objective.²³

Quote 37: “I don’t think you can get away from it [bias]. And obviously it’ll be coloured by the previous experiences you’ve had with that organisation. Whether positively or negatively. And it’s hard to take yourself out of that knowledge and that experience.” (ON Peer)

Quote 38: “... there’s a box at the beginning of the AQA form where you declare your interests. And it is awkward if you’re going to say something critical. But over the years the line that’s come from the Arts Council is make it useful for the organisation, you are a critical friend, be constructive and always back up any criticisms with hard evidence, so one tries to do that.” (AQA Peer)

3. Public assessment

Conducting face to face interviews with members of the public was raised by some non-ticketed organisations as a potential for introducing bias. Whilst it is recognised that this is a limitation for interviewer-led surveys more generally, organisations suggested it may mean that scores are higher than they perhaps should be.

Quote 39: “People were grading everything quite high. I just thought there was a question around that, even though I was often saying to people, please be honest. And yet I almost felt, that people felt compelled to be positive.” (Organisation)

4. Particular survey design elements positively influencing scores

Inherent across all surveying is a tendency for bias, where a sample statistic can over- or underestimate a population parameter. The challenges of bias, such as unrepresentative samples, bias due to measurement error and sampling error and survey bias, are not unique to the CC platform. However, there were particular elements cautioned for bias specific to the CC platform as identified by organisations.

²² Specific challenges around peer recruitment are discussed later in this section.

²³ The definitions of a ‘peer’ or ‘expert’ are discussed later in this section.

Across the focus groups, interviews and open survey responses, consultees frequently reported that the slider²⁴ did not accurately or reproducibly provide a standardised response as there were no numbers on the scale or set points. As a result, there were concerns that public and self-assessors completed it too quickly without much thought on positioning (and with a natural movement to 'swipe' to the right) and when the same response was required, it was difficult to easily reproduce a result.

Quote 40: "I think there's a problem in the positive feedback in the sense that it's quite fun – so encouraging people to take part, I'm using the language like 'it'll be really quick', but then it probably shouldn't be framed like that as they're going slide, slide, slide." (Organisation)

5. Mixed understanding about what exactly is being assessed

Throughout this evaluation, it was apparent that some organisations were assessing (and encouraging the assessment of) the overall experience, whereas others were simply focused on the product in question (i.e. the exhibition, the performance, etc.). This finding indicates that clarity of what should be assessed is needed to ensure standardisation across all assessors and their scores.

Quote 41: "It's all the product. I don't believe it's just the play and everything else is another thing." (Organisation)

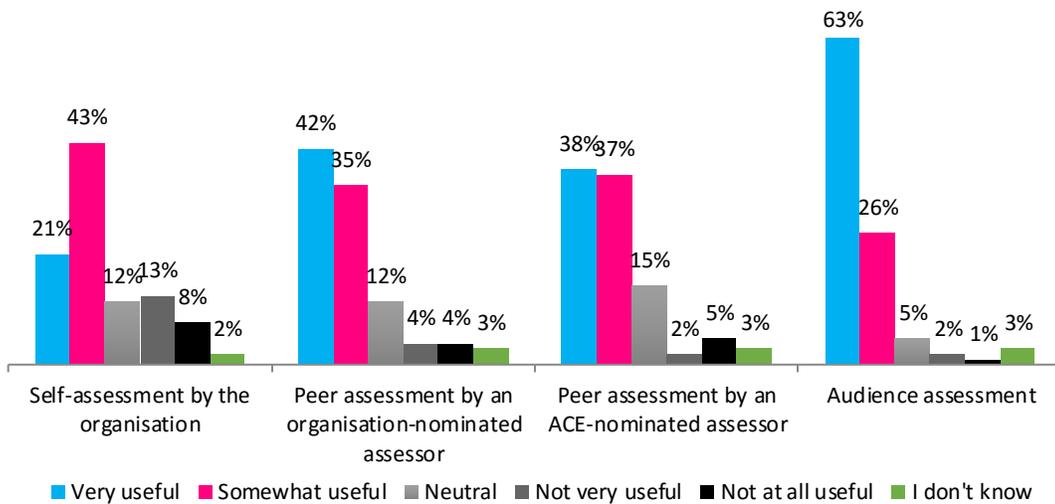
Quote 42: "For me, Quality Metrics is about one piece of work. It's related to that show. That exhibition. That piece of music. A one-off hit." (Organisation)

Major challenges reported with the current peer assessment process

Peer assessments by either ON peers or AQA peers were anticipated to be either very useful or somewhat useful by 77% and 75%, respectively, of organisation survey respondents and as such the second most useful part of the triangulation (approximately mid-way between the audience assessment element and self-assessment). (**Figure 11**)

²⁴ The slider is discussed in more detail in Section 4.2 i.

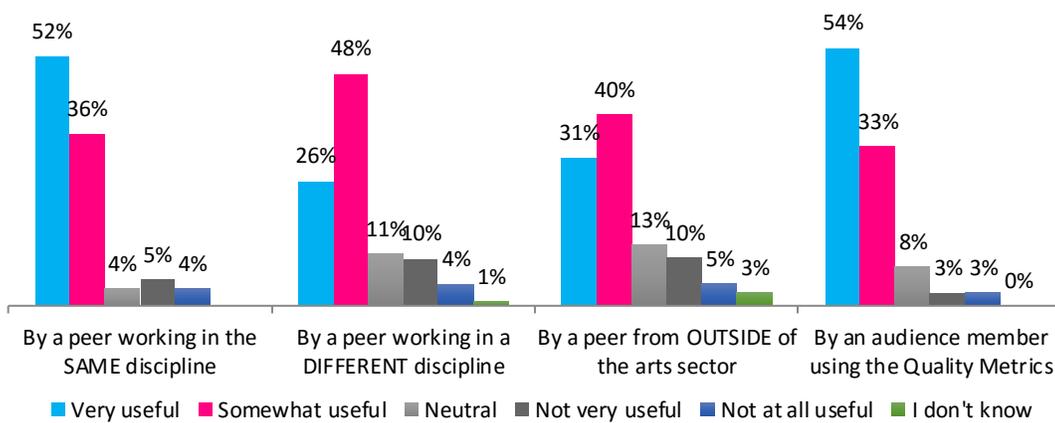
Figure 11: What respondent organisations thought of the triangulation elements



Count = 98

There was even stronger expectation by peers themselves in the potential utility of peer assessment by peers of the same discipline – 88% thought it could be very useful or somewhat useful²⁵. (Figure 12)

Figure 12: Anticipated usefulness of peer and audience assessment to participating organisations



²⁵ Aside from usefulness, the positive impact on peer assessors is discussed further in 4.4.



Count = 112, 112, 111, 112

Despite the potential of peer assessment, the process was frequently referenced as the most challenging aspect of the QM framework for organisations.

Quote 43: "The single most difficult aspect of the process." (Organisation)

Challenges with the peer assessment process were attributed to the following factors.

1. Organisations struggled to recruit peers.

Recruitment difficulties were due to a variety of factors, including the following:

- Both organisation and peer consultees suggested that the location of some organisations created a potential distance barrier, not only for venues in rural areas but also where performances finished late in the evening when public transport options are limited;
- Peers (especially ON peers²⁶) reported that time out of the office or family commitments were barriers. Indeed, a number of ON peers raised the issue of whether the assessment was something they would be permitted to attend by their employer during 'work hours' or whether it had to be conducted on their own time;
- Both organisations and peers reported problems with recruitment emails arriving in SPAM folders or looking like SPAM (i.e., from an unknown source).²⁷ Lack of contact from peers was often attributed to this technical issue, especially by participating organisations who were also ON peers;
- Organisations perceived that there was a potential lack of propensity to attend less 'glamorous' events, with peers choosing instead to select an assessment of a more well-known organisation or experience; and,
- Organisations were clearly confused about which peers to recruit. Some thought they had to only use the provided list of peers nominated by organisations taking part in the trial (separated by geographical areas), others believed they could add their own suggestions to that list, and some did not know a list existed at all.

Quote 44: "That automatic email wasn't enough. I had to chase them and then help them find the email." (Organisation)

Quote 45: "Trying to recruit peer assessors is hard when you've got no profile ..." (Organisation)

²⁶ The time burden for ON peers is discussed further in 4.3 i.

²⁷ Challenges with the dashboard are further discussed in Section 4.2.

Quote 46: “Quite often the Culture Counts emails get lost in the deluge of emails I get, and trying to dig through for the Culture Counts one was difficult, so maybe having a centralised log in would make more sense?” (ON Peer)

Quote 47: “If you come to our theatre from anywhere on the London line, you can’t get train back to show from London so that means an overnight, so we’re paying for an overnight train ticket which becomes expensive and you can see why people won’t want to commit 24 hours. The expense wasn’t the main driver it was the time, whether people could give that time. Of the three events we didn’t get a single person in from that list.” (Organisation)

The impact of these challenges with peer recruitment was stark. When explored in the focus groups, organisations were disappointed that their peer assessment efforts were not reciprocated, leading to an overarching sense of negativity around the QM framework experience. Others gave up recruiting peers entirely, and were therefore unable to benefit from triangulation. Finally, and as discussed above, the lack of ‘official’ peer responses led organisations to rush into finding new replacements, often introducing potential bias due to the selected peers’ familiarity with the organisation. Overall, there were few suggestions related to how the peer recruitment process could be improved. One suggestion was to create a hyper-local pool of peer assessors to resolve the distance barrier, another focused on ‘piggy backing’ on existing behaviours – instilling an awareness of the QM evaluation across the sector so assessors actively volunteer when attending a work.

2. Organisations were expected to cover expenses for peers.

Consultees expressed that they were unaware of the need to budget for this element of the national test phase. There are examples where this caused embarrassment and disagreements within interdepartmental teams as to who was going to reimburse peers or how peers could be incentivised to attend.

Quote 48: “Our events are free so we couldn’t even offer a free ticket. I had to almost kneel down and beg for a free drink for people as it just felt we had to give some kind of incentive to come, at least give them a cup of coffee.” (Organisation)

Quote 49: “I’m having a dispute with the theatre team at the moment about that, because in their view the peers are national, even international has been thrown into the mix, and if you’re talking about judging the work in terms of our artistic objectives then I quite agree, but they say so are you going to pay for it, and I say no I’ve not got budget – so unless you can stump up for hotel and train it’s not going to happen, so it’s like what’s the point of doing it then?” (Organisation)

3. Peers needed greater clarity about what to expect and their role.

Peers commonly reported that they felt poorly briefed before their assessments and had little idea of what it was they were going to be asked following their experience. This caused anxiety in particular for ON peers who typically had less formal experience of assessment than their AQA peer counterparts. It was reported that this lack of clarity resulted in disappointment following attendance, as peers were unable to provide the amount of feedback they would have liked.

Quote 50: “It would have helped to have seen the questions before.” (ON Peer)

Quote 51: “From the first event I went to I took loads of photographs and was writing loads. And then when the survey came, it was oh my goodness, it was really very simple...It felt a bit like, oh, is that it?” (ON Peer)

In addition to the above challenges, there is evidence to indicate that the definition of ‘peer’ may need further exploration and clarification (including expectations of peers’ level of knowledge and experience). ON peers had mixed views in terms of the necessity of being ‘experts’ in the same art form discipline, whereas AQA peers more strongly believed that assessors should have expertise in the same art form they are asked to evaluate.

This finding indicates that further research may be required to examine whether different types of peers and their levels of experience (expert or otherwise) produce greatly different results for the same assessment. Survey results were insufficiently conclusive. (Figure 52 and Figure 53)

Quote 52: “A peer has a certain status attached to it.” (ON Peer)

Quote 53: “Perhaps it might be valuable to ensure the peer reviewers recruited for a particular event or production are a balanced mix, including at least one or two from a similar specialist area?” (ON Peer)

Quote 54: “The framework for the peer assessors – do they need to be experts, or does it not matter if you’re not in the sector or a certain discipline? It’s a real grey area.” (Organisation)

Quote 55: “I just found myself thinking goodness I can’t peer view a pantomime in [name of town] because I’m not equipped to do it.” (ON Peer)

Quote 56: “I’m quite alarmed that organisations might be getting things described as peer reviews that are lay audience reviews. A peer to me is someone who has professional experience and knowledge and expertise in the art form.” (AQA Peer)

Triangulation: more ‘interesting’ than useful at this stage?

Whilst consultees agreed in principle with the concept of triangulation (indeed it was described by one consultee as the ‘USP’ of the whole QM framework) and broadly thought the separate elements of use (Figure 11 and Figure 58), the majority struggled to find value in the resulting data.

Challenges with triangulation were raised because of a lack of confidence in its reliability (due to potential bias introduced at peer recruitment stage and/or through the self-assessment process) or because some organisations had no data at all for comparison from peers (due to problems with recruitment). As previously stated some of these biases are not real but they are perceived and can undermine appreciation of the framework.²⁸

²⁸ See earlier section on ‘Self-Assessment’ under ‘Bias or skew is potentially introduced by particular elements of the framework’ in Section 4.1 ii.

Triangulation places a significant burden on organisations but not equally across the three elements. The majority of organisations reported that audience assessment and peer assessment (both pre- and post-) were very or somewhat difficult to conduct. However, there is a differential impact depending upon size of organisation. Large organisations had greater difficulty than small or medium organisations with peer assessments. Conversely, small and medium organisations had more difficulty with audience assessments (for reasons explained in section 4.3). In regard to self-assessment, the majority of organisations thought it was very or somewhat easy but again there was a differential response according to size. Small organisations, despite their low level of staffing, found self-assessments much easier to conduct than did medium or large organisations (**Figure 35, Figure 36, Figure 37 & Figure 38**). In terms of the value that organisations placed on meeting that administrative burden, they were roughly equally split in thinking whether the time and resources spent on getting sufficient response rates and in conducting the evaluation was appropriate (**Figure 39**).

Peer assessors were more positive in the survey in respect of particular elements of the triangulation, with the majority anticipating peer assessment by an assessor of the same discipline and audience assessment to be 'very useful' to the participating organisation. It is worth noting at the same time that peer assessors were markedly less positive about the utility of peer assessment by a peer of another discipline or from outside the arts sector (**Figure 58**).

For those who had completed full triangulation, the resulting data was regularly referenced as 'interesting' rather than 'useful'. However, it should be noted that more critical views on triangulation may also reflect the level of understanding about how to use the resulting data.

Quote 57: "It's the triangulation and the feedback from self and peers which has made it interesting." (Organisation)

Quote 58: "Because I'm struggling so much with self and peer I'm just going on public so the triangulation means nothing to me." (Organisation)

Quote 59: "...the public side we're happy with, again the deterrent is the peer and the self side of it – I don't know whether the data you get from the peers is worth all the time and effort to get it and how are we using it. I don't see the huge relevance, I know this triangular thing is supposed to be great, and gives us a bit of an overview but if we don't get it is it the end of the world?" (Organisation)

iii. In summary

There is an opportunity for the QM framework to be an effective tool for organisations, as overall, participating organisations could envisage the use of the metrics and were generally of the view that the individual dimensions were of sufficient suitability to be used as measures of quality. However, further refinement and development of the dimensions are required before they will be relevant and applicable for all art forms and cultural disciplines. Organisations are seeking a bespoke, tailored approach that more readily meets their needs – a flexible system where they can select metrics that measure what matters and aligns with their individual artistic objectives. Context is key: organisations and peers need the ability to add narrative to justify scores or explain event assessment selections.



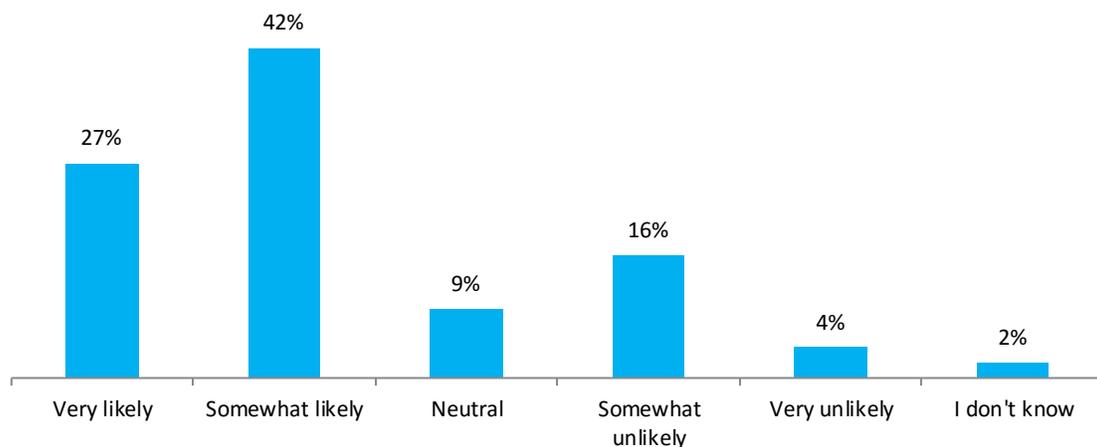
Confidence in the resulting data is limited due to its perceived robustness and there are wider concerns around survey fatigue. Bias has potentially been introduced to the data through a variety of channels including the self-assessment process, multiple interpretations of the metrics, peer recruitment process, and survey design elements such as the slider and face-to-face interviews. Peer recruitment was the biggest challenge for organisations, with some giving up entirely on the peer process. This challenge, coupled with concerns around bias, has led to a lack of perceived value in the otherwise agreeable triangulation process. The definition of a peer is described as a grey area – with mixed views from consultees as to whether peers should be experts in a specific art form or from across the sector. All peers, especially ON peers would have welcomed a clearer briefing in advance of their assessments; at the very least seeing the questions beforehand.

4.2 Suitability and viability of the platform

This section reports on consultees' experiences of using the CC platform; its strengths, weaknesses; and, whether organisations and peers believe it is a practical and efficient part of the framework.

Generally, responding organisations (and focus group attendees) were broadly positive towards using the CC platform when cost was not a consideration. (Figure 13) There were developments and considerations that were highlighted in the focus groups and interviews, which follow.

Figure 13: Likelihood of respondent organisations using the CC platform in future if cost was not a consideration



Count = 102

i. Further development of the platform is needed to gain buy-in

Whilst it was recognised by the majority of consultees that the CC platform was in the early stages of development there were a number of opportunities highlighted that would improve the experience for users with both the front end and back end.

The front end (i.e. the survey software)

The short length of time needed to complete the survey was the most commonly occurring strength of the front-end system. Survey results from the peer assessors and those from participating organisations both confirmed the ease of use of the CC platform. Nearly half (47%) of respondent organisations

somewhat or strongly agreed that the time and resources spent on the CC platform was appropriate (**Figure 39**) and over three-quarters of peer assessors thought the CC platform was very easy to use (55%) or somewhat easy to use (24%). (**Figure 57**).

Consultees reported that self, peer and public assessors found the survey easy and quick to use which, given critical comments raised around the length of other monitoring and evaluation systems such as Audience Finder, was welcomed. The open 'three word' text box was commonly described as a strength, with one consultee suggesting that the qualitative element had proved the most useful part of the resulting data as it was deemed more 'tangible' and accessible to share with fellow staff members than the quantitative information.²⁹

Quote 60: "Because you were doing it face-to-face it was great to say with absolute confidence that it will take you two minutes or less." (Organisation)

Quote 61: "I shouldn't say this perhaps, but takes me five minutes to fill in, it's so quick and easy." (AQA Peer)

Quote 62: "...they gave us three words and we put them into word clouds, and then I took those to my artistic team...they got really excited by them because the graphs didn't work for them." (Organisation)

However, the main weaknesses of the front end system related to two key areas:

1. **The Slider.** As indicated in 4.1 ii, the slider response function was criticised for its likely inaccuracy, difficult reproducibility, and openness to interpretation. From a practical perspective, some users reported they were unable to register a neutral view; that the slider had to be moved from centre. The evidence suggests that these factors led to concerns about the reliability of the resulting data.

Quote 63: "Am I confident in the resulting data? Not at the moment no. The data we're collecting is very grey and biased to the fact that people are using a slider." (Organisation)

Quote 64: "Different people use the slider differently – some people would never go to either end. It's quite open to interpretation." (Organisation)

2. **Accessibility.** In addition to the earlier discussion of language accessibility, the findings indicate that some consultees were concerned about the overall accessibility of completing the survey, especially for those with learning disabilities and visually-impaired or blind users. This finding suggests that further survey design development and bespoke user-testing is needed to ensure that it does not exclude any potential voices.

²⁹ Usefulness and lack of expertise regarding interpretation is discussed further in Section 4.3 iv.



Quote 65: "I would like to see them a bit more accessible, being able to change background and font size so it's easier to see, or have it read it out...I think it's got a long way to go in terms of access." (Organisation)

The back end (survey design and data dashboard)

In terms of strengths, there is some evidence to demonstrate that a number of consultees benefitted from using the CC dashboard, particularly those who did not have another comparable method or who were new to this type of data collection. It was described as simple to use, with a clean interface.

Quote 66: "I've liked using the tool because it's given us a medium for data collection that we didn't have before." (Organisation)

Quote 67: "I thought it actually looked quite clean, I found it relatively easy to use. It was simple but then so was the survey." (Organisation)

However, there were a number of weaknesses which some consultees described had negatively impacted their overall experience of the trial, as well as the resulting data. These weaknesses centred around the simplified design and functionality of the survey, namely the lack of ability to add survey routing, weighting, customisation, mass email mail-out (as opposed to one at a time) and the repeated work required to produce templates.

Quote 68: "In terms of being a fair trial I think the lack of flexibility of the tool, the lack of rigorousness of the tool, in terms of reporting, in terms of the way you ask the questions - it meant that the trial becomes meaningless. It's all very well saying you're developing the tool but the tool is crucial." (Organisation)

Quote 69: "The simplicity came at a cost. Of actually then collecting usable and useful data." (Organisation)

Quote 70: "That repetition of survey design feels very laborious. I kind of feel it should automate – here's your peer, here's your self, here's your public." (Organisation)

Quote 71: "...you couldn't add any demographic questions – we wanted to ask ethnicity. And that's one of the ones that ACE is desperate to know and we have to give them, so how come they couldn't add that in." (Organisation)

Whilst consultees praised the support provided by the CC team,³⁰ some struggled to get to grips with technical aspects of the survey templates, which required more time to be spent on the project than anticipated and impacted their remaining workload.

³⁰ This is discussed in Section 4.2 ii.



Quote 72: "Time was a real challenge. Some of the set up things, I set it up incorrectly, then you have to untangle it." (Organisation)

Quote 73: "The time was set in Australian time so we couldn't set up our surveys. We had difficulties with setting up individual surveys, setting up three surveys each time. We just had a lot of technical issues." (Organisation)

However, one of the most frequently cited weaknesses was the reporting functionality and the lack of automatic insight. Consultees who were interested in benchmarking³¹ were unable to access comparative data (citing the reporting option provided by Audience Finder or Visitor Verdict as a 'good' example) and some who wanted to cross-tabulate findings with demographic data were unable to do so. The opportunity did exist in the form of csv files that could be downloaded (for which information was available online and CC would have provided support too.) but which clearly was not readily found or used by the said consultees. In addition, consultees had expected the system to automatically create sophisticated reports and without these they were left unsure of a) how to use the data and b) how to communicate it in an informative and interesting way to their colleagues. There is evidence to suggest that this limitation particularly negatively impacted on those who were less experienced with manipulating data (particularly in terms of the additional unexpected time spent on the national test phase) and on the perceived usefulness of the overall national test phase.³²

Quote 74: "I found the reports hideous to decipher...I took it as it stands to my artistic director and said what does this tell you, and she just pushed it away and said 'absolutely nothing'." (Organisation)

Quote 75: "...it would be good to have that in terms of individual art forms or an overall aggregate, but just to have that measure of what the benchmark is and the national average is. I think that would be useful for us and ACE." (Organisation)

Quote 76: "I thought there was going to be more support with processing the data. The data's there on screen, but then what do I do with that and how do I communicate it with the rest of my organisation? I've created a report and that's been time consuming." (Organisation)

ii. The Culture Counts support team were highly regarded

Overwhelmingly, the consultees were satisfied with the support they received from the CC team who were reported to help wherever possible. They were particularly praised for supporting organisations in the early stages of the project when any issues arose with how to use the technology. Survey results from participating organisations very clearly showed that when addressing problems with the CC

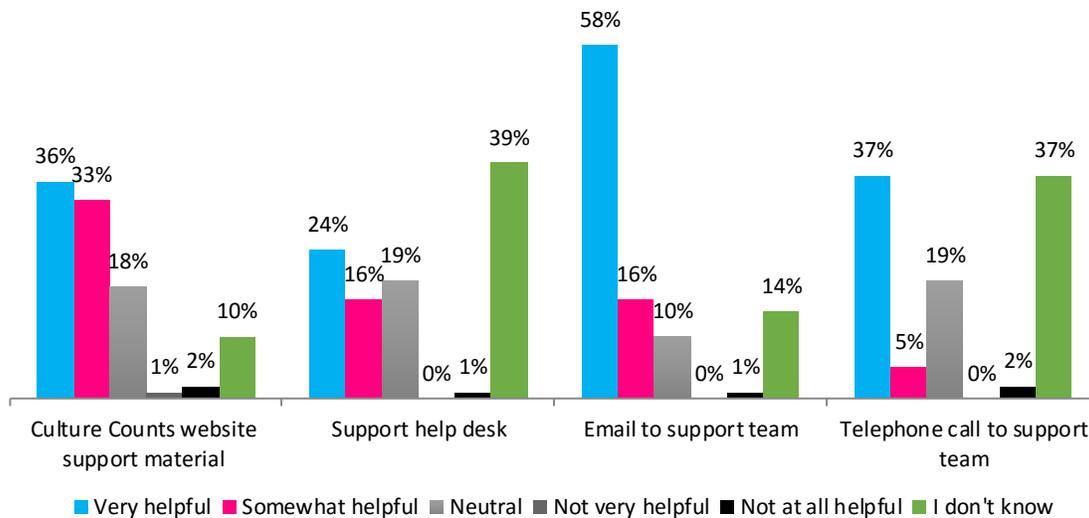
³¹ Attitudes towards benchmarking and sharing data is discussed further in Section 4.4.

³² Usefulness of resulting data is discussed further in Section 4.3 iv.

platform, the support team was prompt in their response (79% rated it as ‘very positive’) and effective (67% rated the response as ‘very positive’). These findings were based on the assessment of the 44% of respondents that had sought help (Figure 41 & Figure 42). More generally in regards to the different means of receiving support, the majority of respondents who had used the service had found both emailing or telephoning to be ‘very helpful’ and the website support material to be ‘very helpful’ or ‘somewhat helpful’ (Figure 14) Figure 43.

Quote 77: “I thought that the team running it did try hard to help.” (Organisation)

Figure 14: Helpfulness of each medium of support available for CC platform



Count = 98

iii. Connectivity and kit

A range of organisations undertaking face-to-face interviews faced a number of challenges with the Wi-Fi connection required for survey completion. This difficulty was not only experienced by those in rural locations and outdoor arts events, but also by those in modern buildings. Indeed, lack of connectivity was cited by non-participating organisations as a barrier to participation in the national test phase. For those participating organisations who experienced connectivity problems they reportedly swapped to paper-based surveys and manually inputted the data instead, which added an extra burden in terms of time. This finding suggests that an off-line solution may have encouraged more non-participating organisations to take part and also created a smoother experience for those with Wi-Fi problems.



Quote 78: “It should just keep going...if you had an iPad and had an app, you can use it and then when it connects to Wi-Fi it just downloads all your stuff instead of just having a constant connection. Even in our building there are dead zones.” (Organisation)

Non-ticketed organisations commonly reported that they had to purchase iPads in order to conduct the face-to-face interviews with visitors. This requirement added an extra level of complexity and cost, as the majority had to specifically buy-in kit (often unexpected). As availability of kit was a cited barrier for non-participating organisations too, the findings indicate that alternative solutions may need to be found for organisations who have little budget, whether that is delivered through a system of loaning tablets (as suggested by one consultee) or supporting organisations to set up alternative data capture arrangements such as manually collecting emails and circulating the survey online post-visit.

Quote 79: “We don’t own any iPads and I’m probably not going to buy any iPad, because they’re not relevant to what we do. We’re a small organisation. We don’t have those funds available. And I found that difficult that there was no other way to engage with the audience.” (Organisation)

Quote 80: “...we didn’t have the hardware or the budget to acquire the hardware...” (Non-participating organisation)

iv. In summary

The ease with which the survey could be completed and the use of the CC dashboard by organisations with no previous experience of similar tools, were both seen as strengths of the system. This was alongside overwhelming feedback through all elements of the consultation that the support provided by CC both in terms of quality and timeliness was highly satisfactory.

However, to improve the viability and suitability of the CC platform a number of early developments would be needed to both the front-end (survey software) and back-end (survey design and dashboard). Most notably to the ‘slider’ in the survey (to remove a perceived bias) and by ensuring that the approach is inclusive through improving accessibility of the user experience. Whereas in regard to the back-end, the survey functionality for designing and disseminating the CC surveys was considered to be over-simplified which often resulted in difficulties using the technology and software, with a consequent increased workload. What was most frequently cited though was the lack of sophisticated reporting functionality to produce more useful analyses or reports. At a practical level, it is not possible to ensure continuous connectivity for all organisations in the portfolio and so an offline solution needs to be found otherwise translation to paper-based mode adds to the administration burden.

4.3 Capacity, usefulness and alignment

This section reports on consultees' experience of administering the QM framework and considers whether the resulting data is insightful and how organisations envisage using it. It also evaluates whether there are merits in exploring opportunities for alignment with existing systems such as Audience Finder and the Artistic Quality Assessments.

i. Capacity burdens

Consultees commonly reported that they had found the process unexpectedly burdensome – especially participating in training, setting up surveys, peer recruitment, collecting the data, and interpreting and reporting on the data.³³ (Figure 15 and Figure 16) Lack of capacity for fieldwork was particularly problematic for non-ticketed organisations or those without the capability to use box office data to email out surveys. Although some organisations chose to recruit volunteers to undertake the fieldwork,³⁴ this required additional time to be allocated for training.

Smaller and medium sized organisations struggled generally in terms of finding time and staff to dedicate to the project. (Figure 38 and Figure 37) In addition, a lack of capacity was cited by all non-participating organisations consulted as a major barrier to participation, including a perceived inability to logistically run both Audience Finder and the QM framework concurrently.³⁵

Quote 81: "I don't have time to use the training videos." (Organisation)

Quote 82: "It's about capacity again. So if we were to do this for every exhibition and event, we'd have to recruit a new person! And for us that's just not feasible. Another benefit of Audience Finder is that it runs continuously and it aggregates everything. To me you're not going to set up a survey each time." (Organisation)

Quote 83: "We just didn't feel we had the capacity within the timeframe...and the fact that it's so close to Audience Finder which has taken additional resources that only go so far." (Non-participating organisation)

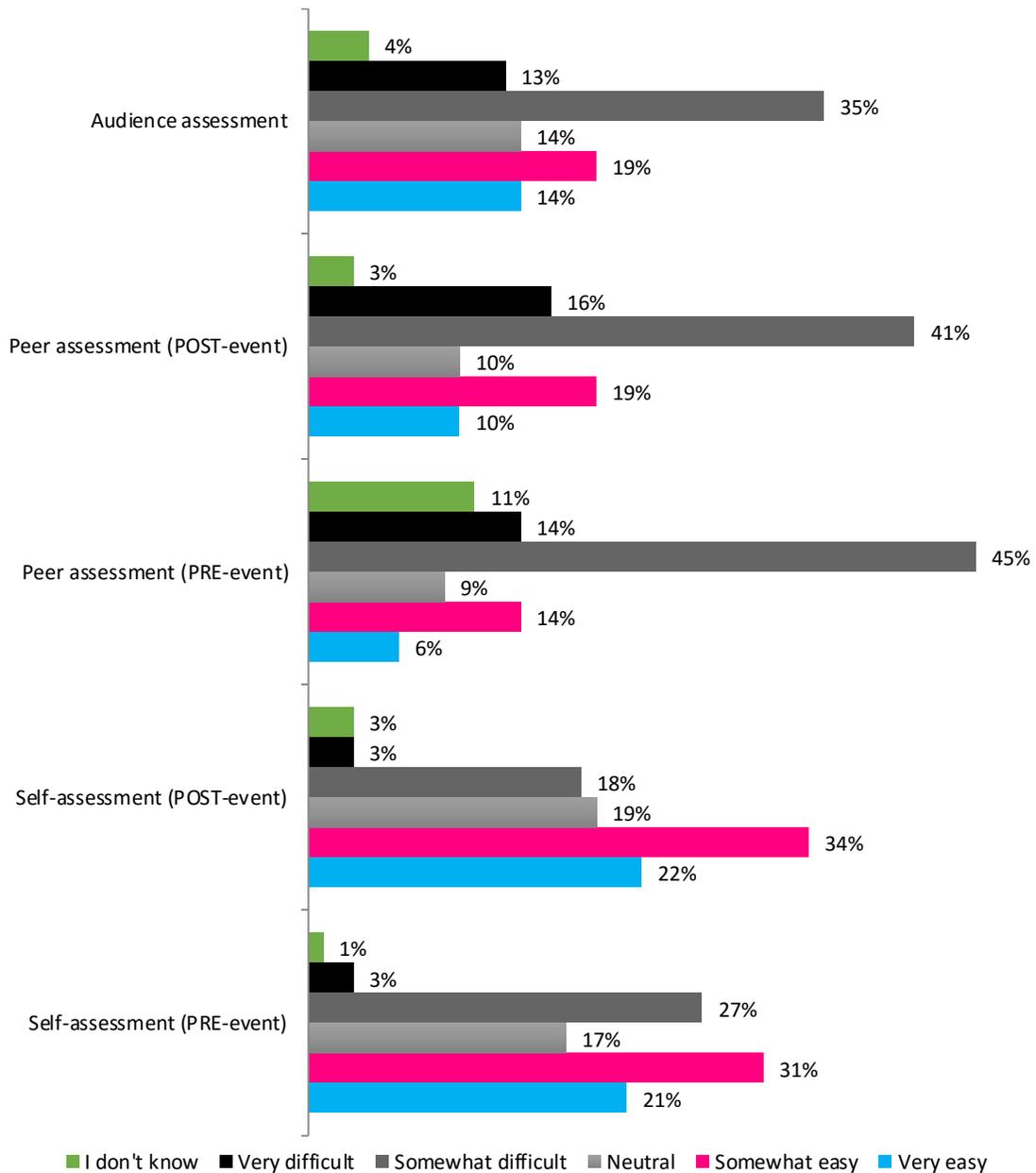
Quote 84: "I haven't used it that much. And that's partly because there's no one on our team dedicated to this work, it's nobody's job to spend a lot of time doing this. I think the data requires quite a lot of analysis and interpretation and we just don't have the capacity to spend that time with it." (Organisation)

³³ Details of the survey results about the administration burden can be found in Section 4.1 li under triangulation.

³⁴ The benefits in terms of CPD for volunteer fieldworkers/front of house staff is discussed further in Section 4.4.

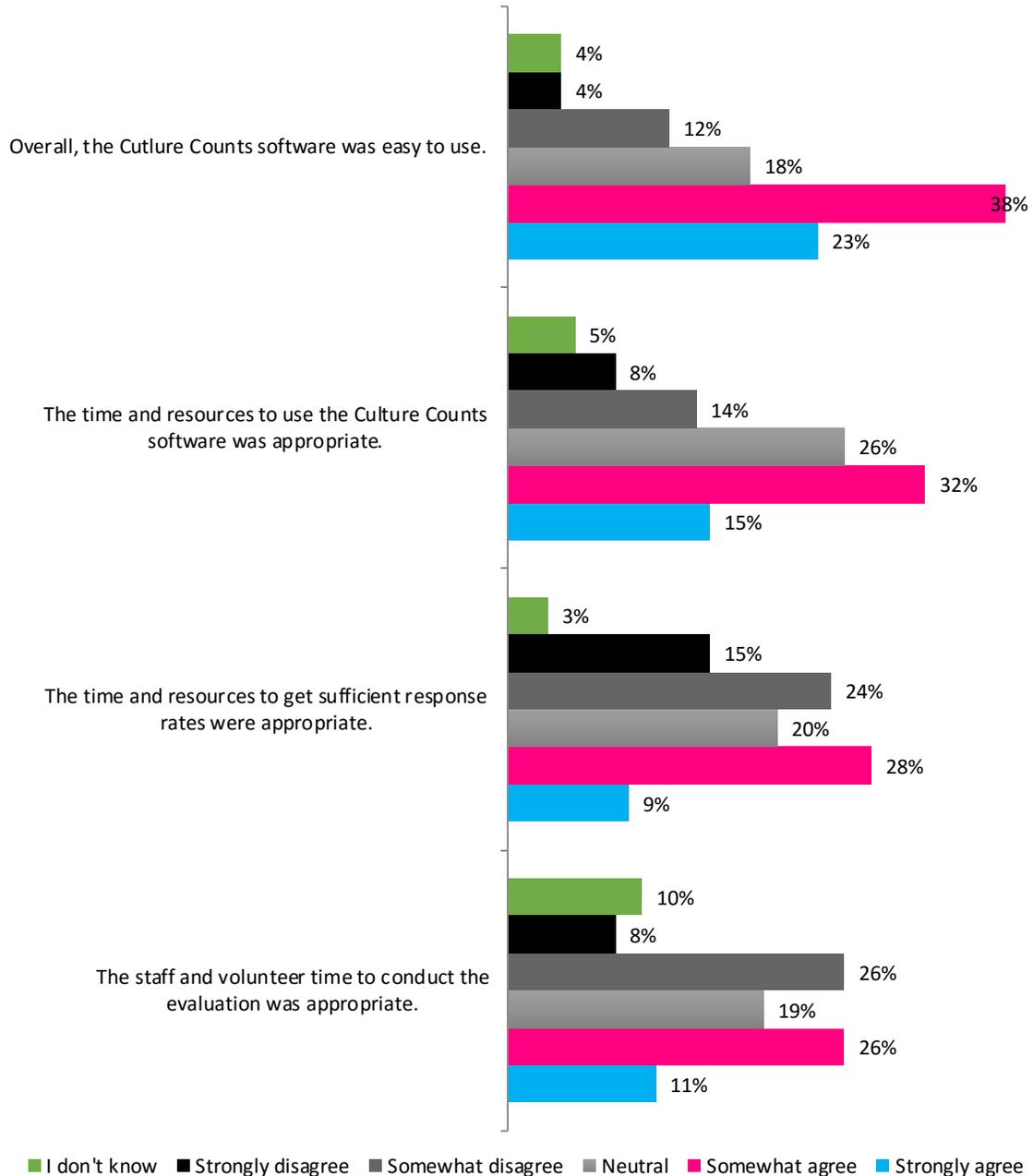
³⁵ This is further discussed in Section 4.3 iii.

Figure 15: The administrative burden for coordinating and/or conducting each element of the triangulation



Count = 98

Figure 16: Further insight into what respondent organisations thought of the process



Count = 98

In addition, the findings suggest that ON peers were more likely to report time pressures than AQA peers – not altogether surprising, given that AQA peers are paid an annual fee, have been through a rigorous recruitment process and training, and are expected (and presumably willing) to make time to conduct an assessment. As indicated earlier, the evidence suggests that there may be additional complexities and time barriers for ON peers to attend assessments because of their existing workload³⁶.

Quote 85: “The QM is utterly un-burdensome. All the work if you’re an AQA assessor is at least a day’s work. Going to an event and writing it up, and that’s where you earn your fee, the QM is only five minutes on the back of that.” (AQA Peer)

Quote 86: “Taking part as an assessor made me realise how much time it takes up outside of my allocated working hours. Travel to and from shows took up a lot of time, which then impacted negatively on my personal working hours.” (ON Peer)

Quote 87: “If there was an evening event by the time I’d driven from [name of town] I wouldn’t get back until midnight and I would have to be up early in the morning for work.” (ON Peer)

Quote 88: “We also have a huge time commitment producing our own work, so there were a lot of evenings where it wasn’t possible to go out to see work as we had our own on here, so it’s just human capacity.” (ON Peer)

ii. **Greater levels of expertise and training are needed**

As outlined above, consultees reported that they spent more time than expected trying to make sense of the resulting data. This in itself exposes the varying levels of expertise and experience with analysing, interpreting and using data more strategically. Reflecting the wider challenge of ‘data shy’ organisations in the sector, it also highlights a specific training need that may enable participants in the national test phase to more readily benefit from the QM framework evaluation. One consultee also highlighted the legacy and sustainability challenges that exist due to staff turnover, and the need for the ongoing training solutions to ensure that new staff can quickly get to grips with the system.

Quote 89: “I think one of the issues with the sector is we have a turnover of staff, so there’s the continually retraining of staff. If you’re going to stay standardised training should be delivered by Arts Council.” (Organisation)

Quote 90: “I know organisations who don’t have an expert – if you gave them the data they’d not have the confidence to do anything with it.” (Organisations)

³⁶ There are a number of ON peer CPD benefits to note despite barriers to time, these are discussed further in Section 4.4.

iii. Audience Finder, AQA and Museum Accreditation

A lack of time and resource to run the QM framework and Audience Finder simultaneously was a challenge for the majority of organisations taking part (most of whom chose to either run one or the other during the national test phase period). Whilst this was predominantly attributed to survey fatigue as described earlier, non-ticketed organisations elected to run only one survey due to the fieldwork capacity challenges cited above. This indicatively impacted on the data collected for Audience Finder and other individual evaluation and monitoring systems across the national test phase period, with these essentially being displaced by the framework.

Quote 91: “I could only choose to do one given my capacity. I chose not to do Audience Finder and that came as a surprise to my Relationship Manager.” (Organisation)

Quote 92: “...at the time we were already doing Visitor Verdict, Visitor Finder and internal evaluation, and in terms of the people delivering that, it was a huge amount of fatigue.” (Organisation)

Consultees expressed mixed views about how to resolve the above challenges, and whether existing evaluation systems like Audience Finder could be more readily aligned. The majority described the tools as ‘different beasts’ with different objectives, and organisations who had tried to combine the two were not confident it had worked well. This suggests that whilst there is a capacity burden in running two separate surveys, there are also concerns around both the practicalities and questionable methodology of simply merging the two.

Quote 93: “It’s utter nonsense to have competing, parallel systems. Quite clearly it’s not sensible.” (Organisation)

Quote 94: “It might be better to integrate some of the metrics into Audience Finder.” (Organisation)

Quote 95: “I don’t think they fit together at all. Audience Finder is about profiling.” (Organisation)

Quote 96: “It’s a bit tricky if you’re going to try and merge things together which are not supposed to be that don’t fit – Quality Metrics we can do in a space and the other things are an exit survey.” (Organisation)

As described earlier, AQA peers were concerned that the QM survey lacked any opportunity to add justification and narrative, in comparison to the existing AQA format³⁷. Whilst feedback from consultees on the current AQA approach and resulting reports was mixed – two particularly critical of their value –

³⁷ <http://www.artscouncil.org.uk/national-portfolio-organisations/artistic-and-quality-assessment>

AQA peers would be willing to complete the QM dimensions, but only as an addition to (i.e. not replacing) the existing AQA process.

Quote 97: “It’s just a totally different exercise. AQA is focussed on detail, evidence gathering and really if it’s done well provides organisations with quite specific feedback on what they’re doing well in terms of performance quality. This [Quality Metrics Framework] is so general and broad brush with no real opportunity to explain or illustrate or to go into any detail. It’s of very limited value from that point of view.” (AQA Peer)

Quote 98: “It feels less arbitrary or random than the Artistic Assessments that they send people out to, because we have no say in who that assessor is. And we’ve had problems with that...” (Organisation)

It should be noted that two consultees within the evaluation queried whether there would be merit in exploring the benefits of the QM framework with the Museum Accreditation³⁸ process. However, no other alignment opportunities were suggested.

Quote 99: “It’s interesting to know how this might feed into accreditation and if it has any value to it.” (AQA Peer)

Quote 100: “I don’t know if it applies to theatre but in M&G we have accreditation which is a huge piece of work, and it’s how this that fits with that, and whether it meshes together?” (ON Peer)

iv. Platform Providers

Consultation with platform providers (both alternative and existing) expressed strong support for the initiative setting out to measure quality and for the development of the open source QM framework, however, views on the delivery platform were conflicting. The triangulation approach to evaluation (of self, peer and audience assessments) was applauded by all platform providers, and the development of the framework was deemed both innovative and seemingly effective. The provision of new data was welcomed as an opportunity for the sector.

However, concerns were raised about tying QM to one particular platform provider at the exclusion of others. The issues included:

- data ownership and privacy issues (also related to concerns expressed by organisations), including a platform provider having preferential access to the accumulated data;
- the impact of organisations having the onus of managing multiple data platforms (including having a detrimental effect on the use of their own platforms); and,

³⁸ <http://www.artscouncil.org.uk/supporting-museums/accreditation-scheme-0>



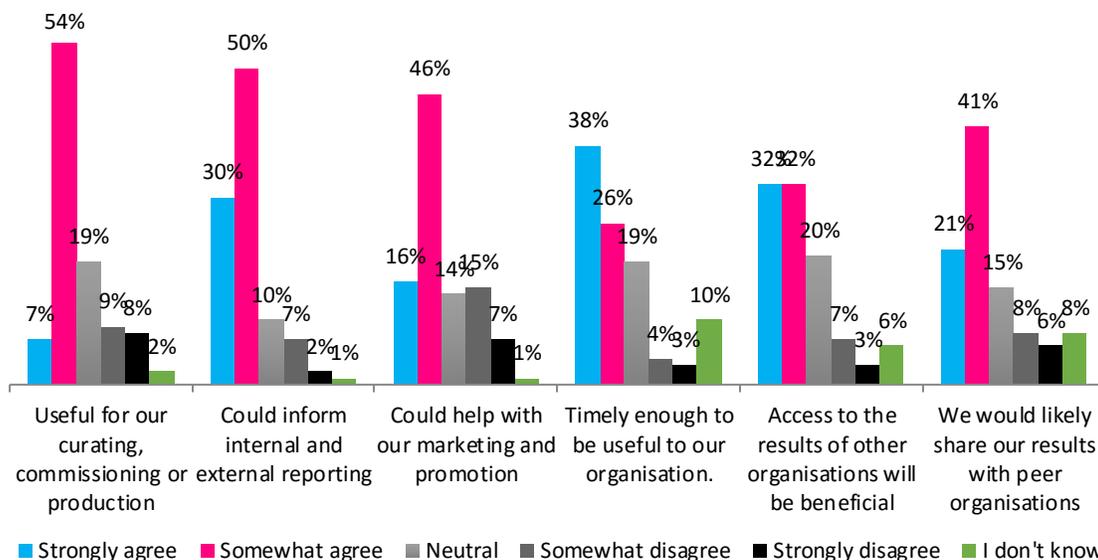
- a risk of the cost of a proprietary platform increasing at a later date.

The enthusiasm for QM by platform providers was expressed by their common interest in harmonising or integrating the QM framework with their existing platforms.

v. Usefulness

The survey results indicated that the timeliness of the process, the potential to fulfil internal and external reporting requirements and access to other organisations data were all seen as potentially useful outcomes of using the QM framework. It was also seen as being of use for other activities, such as for curating, commissioning or producing and for helping marketing and promotion. Organisations also indicated that in principle they were willing to share results with peer organisation to some extent (Figure 17).³⁹

Figure 17: In what way could the QM framework help the respondent organisations



Count = 98

However, through the interviews and focus groups, consultees were unclear about how to strategically use the resulting data from the QM framework and even 'data driven' organisations struggled to see how it could be used beyond reporting upon activity. Even though most organisations had yet to complete the national phase of the trial, there was little evidence (including at the focus groups) of how

³⁹ Attitudes towards benchmarking and sharing data are discussed further in Section 4.4.

organisations intended to feed the data into strategic discussions around programming and self-reflection.⁴⁰ However, it should be noted that though this was an extensive trial, it only ran for a short period, with the majority of organisations wholly new to the concept.

A number of consultees described that they collect data about quality through their own surveys, qualitative methods or Audience Finder – and as a result there were mixed views as to whether the framework had added anything new. Once again, there was a sense of the data being ‘interesting’ but uncertainty as to whether it was ‘useful’; that it provides ‘nuggets’ of information but is more valuable when added to a range of other (often qualitative) data to tell a story. This could be indicative of the challenges that organisations faced with the reporting functionality of the data dashboard, however further support and training from the outset to explain how to use the results was identified by organisations as a missing element of the national test phase that would have been helpful.

Quote 101: “I see it as something that I use to report back on rather than inform what’s going to happen.” (Organisation)

Quote 102: “I would see this as a luxury rather than essential. As we collect data in many other ways, this adds to what we do, rather provides us with all our information.” (Organisation)

Quote 103: “It won’t change the way we do anything. It’s not going to feedback into our programming.” (Organisation)

Quote 104: “It didn’t really tell us anything we didn’t really know.” (Organisation)

Quote 105: “...it has to be used alongside other things that give more context and detail.” (ON Peer)

Whilst the majority of consultees reported that participation with the national test phase had not changed their internal use of data and evaluation in any way, positively there is some evidence to suggest that it has stimulated conversations within organisations about data and intelligence. For example, shifting the evaluation ‘role’ of staff away from simply that of marketing teams.⁴¹

Quote 106: “Up until now it was a thing that marketing did. The acknowledgement that data collection is very much about how we as an organisation need to behave going forward, this pushed an acceptance I was struggling with...this brought something to the surface that wasn’t acknowledged before.” (Organisation)

⁴⁰ Described as a ‘vital litmus test’ at <http://www.qualitymetricsnationaltest.co.uk/blogcontent/2016/1/28/valuing-culture>.

⁴¹ Other unintended outcomes related to organisational learning are discussed further in Section 4.4.



vi. In summary

The national test phase of the QM framework has clearly revealed a capacity issue within the participating NPO/MPM group – a twofold challenge that is the result of the administrative burden placed upon organisations; and the human capital/skills gap of a sector that has further to go in developing sector-wide data competence. The skills gap undermines the framework, contributing to the administrative burden, which is not offset by useful data, as again the lack of expertise and sophisticated reporting mechanisms prevent thorough interpretation and application. Consequently, the potential usefulness of the framework is not actualised across all organisations, even those which are 'data driven'. Although, many participating organisations anticipated the data had the potential to be useful, particularly in regard to informing the commissioning or presenting of work, the marketing of an organisation's programme and in fulfilling reporting requirements of a funder.

The challenge of trying to operate both Audience Finder and the QM framework was significant for the majority of organisations and resulted in most either running one or the other during the national test phase period. The streamlining of such processes would be of benefit but views were very mixed as to how or even whether this was readily possible. Most organisations saw the two tools as quite different, each with differing objectives and even though the CC survey was praised for being quick to answer, it can not be a simple matter of leveraging the CC survey into the Audience Finder survey. Streamlining with the AQA process has some potential, as AQA peers felt they could readily add the QM dimensions into the existing AQA format, however, the dimensions by themselves would be insufficiently detailed to replace AQA.

4.4 Overall attitude and propensity for future adoption

This section reports on the attitude of organisations towards the QM framework, exploring their motivations for participating, key concerns and propensity for future adoption. It also examines organisations’ willingness and readiness for sharing data and highlights potential areas that the Arts Council may wish to address.

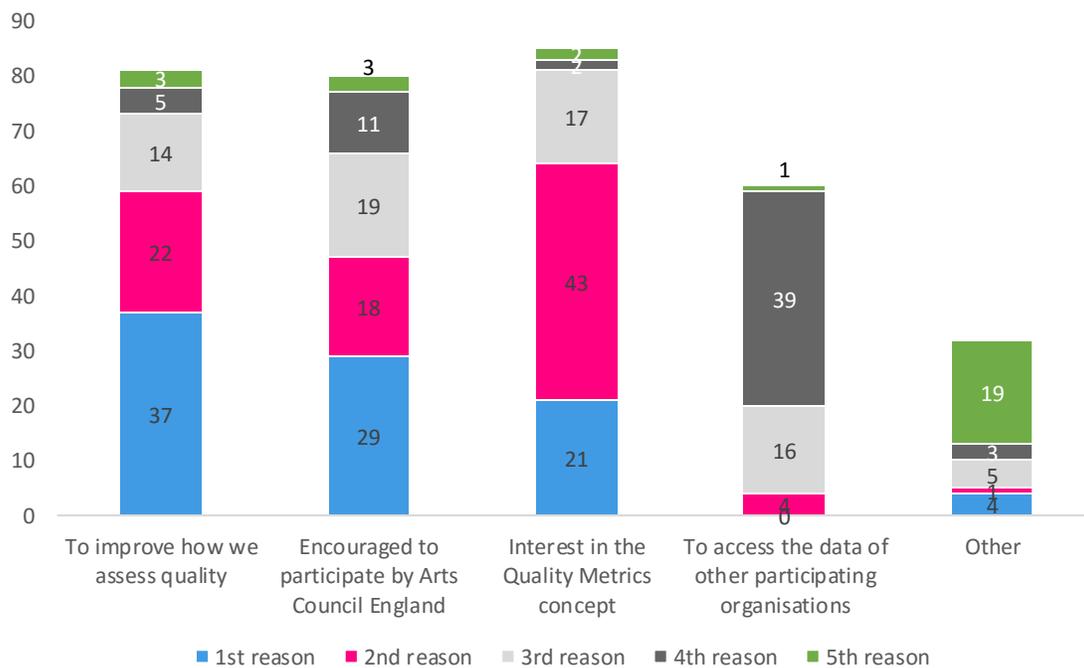
Overall the sector supports measuring quality

There was clear support in the consultation for improving quality assessment and a specific interest in the general approach of the QM framework, notwithstanding a recognition of the innate difficulty in defining quality and of achieving cross-discipline applicability of a single set of metrics. (Figure 18)

Quote 107: “It’s clunky right now but it’s a really great privilege to be a part of a process that’s looking at more quality of things than just quantity and figures and box ticking...it’s not really fit for purpose for everything we’re talking about. But it’s progress in my opinion.” (Organisation)

Quote 108: “Creating established quality measures that can be used widely will help with assessment and impact, so really welcome this development.” (Organisation)

Figure 18: Motivation of respondent organisations to participate in national test phase (frequency)



Count = 81, 80, 85, 60, 32

Nearly three-quarters of responding organisations placed improving their own assessment of quality as their primary or secondary motivation for taking part in the national test phase, with over three-quarters of respondents also citing interest in the QM concept itself as primary or secondary motivations. Indeed, the majority of non-participating organisations did not take part due to practical challenges as described earlier, rather than any negative criticism towards the approach.⁴²

Quote 109: “We agreed in theory that yes it was something we wanted to be part of. That it’s something worthwhile ... investigating if it’s possible to create meaningful national benchmarks for the sector.” (Non-participating organisation)

Quote 110: “We’re a new NPO and we’ve just started with Audience Finder ... we were genuinely interested to see if it was different ... if it would give us anything new. We were interested in how it would differ from what Audience Finder was giving us and we were happy to kind of contribute to the trial.” (Organisation)

Quote 111: “...for us it was about having another dimension of information we could use to reflect on that wasn’t a statistical figure. It was a chance to look at the artistic product and if our assumption about experience was actually real.” (Organisation)

There is scepticism that a single framework can achieve true quality assessment across disciplines and some caution whether it is desirable to render quality into a few simple statements. Nonetheless, some form of assessment is seen as desirable, and as described earlier, a number of organisations reported that simply participating in the national test phase had stimulated additional conversations within their organisations about quality and the use of data. There is also some evidence of unintended learning outcomes, especially for front of house teams who have reportedly gained confidence in survey skills and audience engagement as a result of participation. For these particular organisations, it has helped them improve overall customer experience and provided additional CPD opportunities for their staff. Similarly, a number of ON peers reported that they too had professionally benefitted by taking part in the process, by having dedicated time to take a break from the office and see other programmes, organisations and ideas. This indicates the potential impact that a new evaluation model can have on participating organisations, and is another potential USP to note for the framework in terms of messaging.

Quote 112: “One of the interesting things is that it’s giving the team empowerment to talk to people as a tool and I think that’s when it starts being really useful.” (Organisation)

Quote 113: “... it would be a nice piece of customer care. The value for the relationship [with audiences] is as important if not more important as the value of the data for me.” (Organisation)

⁴² Only two non-participating organisations expressed criticism of the overall approach as a key motivator for not taking part in the national test phase.

Quote 114: “I thought it was a wonderful opportunity to see what other arts organisations are doing and to see how our programme sits.” (ON Peer)

Quote 115: “I found it a really valuable opportunity because often we intend to get out and see events, but that’s often through personal networks and actually it was nice to see outside your network. I think it’s really important to build cohesion in the network by actually seeing the artwork not just at a reception or event.” (ON Peer)

Greater buy-in is needed

Indicatively there were different attitudes to the QM framework depending on whether the participating organisation QM lead had a programming/curation focus or a marketing role. Whilst this reflects the wider and often reported ‘rub’ between marketers (and the product function of marketing) versus programming/curation teams, it demonstrates the buy-in that is needed from both teams if the ultimate intention⁴³ is to use the resulting data strategically and for self-reflection, as opposed to simply reporting on activity. Some consultees reported limited buy-in across the organisation, whether from a marketing or curating background, which at times had hindered the self-assessment process – for example, achieving requested deadlines for completion.

Quote 116: “It should not be me sitting here. It should be the programmers sitting here. Until the programmers are sitting here this is a problem.” (Organisation)

Quote 117: “There is no way my artistic team would take outside influence in the creation of their art, it’s just not the way we work.” (Organisation)

Quote 118: “At the moment it feels like yeah it’s an experiment, there’s a bit of ‘I’ve run out of time, can I extend the deadline when it closes?’ for them [self-assessors] to fill it in, I think they understand it in concept but the reality of fitting that into everything else, that’s not quite happening so it needs more buy in.” (Organisation)

A lack of shared ownership within individual organisations indicatively contributed towards the various challenges experienced with buy-in. However, there is also evidence to indicate that a lack of understanding about the development of the QM approach and its background (especially being ‘sector-led’) prior to the national test phase potentially hindered this too. For example, some organisations believed this was the first time the QM framework had ever been tested,⁴⁴ and a number of non-participating organisations reported a lack of motivation for participating because they had not been involved in shaping the QM dimensions.

⁴³ This is further discussed in Section 4.4 under ‘Clearer articulation of purpose is essential’.

⁴⁴ It is recognised that previous trials were referenced in the national test phase expression of interest and associated printed and online materials from Arts Council England and CC.

Quote 119: “... we were essentially asked to do a survey that we weren’t driving. If someone else tells you something is important and it’s not already on your list of things it rarely rises to the point where you’re going to invest in it.” (Non-participating organisation)

Quote 120: “I heard a rumour that this was from Australia? And is there any data from those organisations to see?” (Organisation)

Quote 121: “Please open a dialogue with venues that are good with data and audience surveys to ask their opinion on new initiatives...we have some good ideas.” (Organisation)

Clearer articulation of purpose is essential

As outlined earlier, there was a common lack of understanding about how to use the results from the QM framework evaluation. However, consultees were also highly suspicious about how the Arts Council intend to use the data. Whilst consultees hoped that the data would be used for high level lobbying and advocating for the sector, there was little appreciation of how the Arts Council would practically use it ‘on the ground’ with organisations. It is evident that a greater clarity of purpose and stronger messaging around the Arts Council’s intentions may have led to increased buy-in of the national test phase and in turn, may have more positively influenced propensity for future adoption. The two most commonly occurring points raised about purpose were:

- consultees were unclear about how the Arts Council intend to use the data in decision-making processes, particularly with regards to future funding, without any opportunity for justification from organisations (reflecting the issues raised about lack of context and the extent to which the QM dimensions are appropriate, as explored earlier); and
- lack of clarity about whether the evaluation is primarily about critical reflection or benchmarking. Consultees were apprehensive about the latter due to the potential misunderstanding or misuse of data that could occur when comparisons are being made between organisations (again, given the lack of ability to add any context in terms of organisation type, product and artistic objectives which may lead to ‘blunt’ or ‘misinformed’ comparisons) as well as confidentiality issues.

Quote 122: “... it could be used as a stick to beat and something that you have to achieve and a hoop to jump through rather than actually help raising the quality of arts.” (Organisation)

Quote 123: “... what worries me is that it’s useful to the Arts Council to start being able to say we think that X’s work is less well rated than X’s work so we’re going to give them more funding to X because they’re creating better work ... that’s the fear.” (Organisation)

Quote 124: “I think we need to have permission to pull up some bad stuff. Or to articulate what hasn’t worked. Because we are in a relationship with the Arts Council, their funding is important and I think they need to be up front as to why we’re doing this.” (Organisation)

In addition, further scrutiny of organisations’ motivations for participation mirrored such suspicions. For example, some participated in the national test phase because they were wary about what may be

introduced and therefore wanted to have a 'seat at the table'. Furthermore, some consultees (including non-participating organisations) suggested that being 'strongly encouraged' by their Relationship Managers to take part had also contributed to a sense of suspicion around a compulsory system ultimately being introduced.

Quote 125: "We wanted to get in front of the race before it came part of it [monitoring and assessment]." (Organisation)

Quote 126: "I've got an NPO application to write, and I want to know what I'm letting myself in for. Because as sure as eggs are eggs there [will] be something in there from this." (Organisation)

Quote 127: "... we were encouraged by our Relationship Manager at the Arts Council to apply, quite strongly." (Organisation)

Mixed views on sharing⁴⁵

Whilst organisations expressed support for data sharing as part of the growing 'big data' culture, some were concerned about how benchmarked data could potentially be misused as stated above. Across the board there were differing opinions of the perceived usefulness of sharing quality data – for example, some reported it could turn the QM framework into a more useful tool overall (especially if presented in a similar way to Audience Finder in terms of discipline/art form, regional and national comparisons). However, consultees found it difficult to articulate how it would be relevant or useful because of the range of different organisations within the portfolio, all of whom have varying organisational or artistic objectives.⁴⁶ Three consultees also queried the commercial implications of sharing data for artists, tours and sponsors – especially if the QM dimension scores were lower than expected and available within the public realm. Overall, this highlights that further exploration of the expectations around benchmarking, its implications for those involved, and communication of its potential uses would need to be considered.

Quote 128: "Benchmarking is really important for us. You need to know for yourself and for reporting to trustees and local stakeholders, getting reliable benchmarking is really important." (Organisation)

Quote 129: "It's all about us individually how we are striving to improve, not how we are comparatively." (Organisation)

⁴⁵ Organisations were able to share data in a basic format through the dashboard however there was no automatic 'benchmarking' functionality as used in similar systems referenced by consultees e.g. Audience Finder. There was little evidence of organisations sharing data through the dashboard as part of the national test phase, although many had taken part in the CCLearning days to debate and discuss aspects of the project and their own findings.

⁴⁶ This may reflect some of the findings related to lack of understanding and expertise about how to use the data discussed earlier in Section 4.3 iv.



Quote 130: "I think there has to be a consideration around this programme, how it affects artists, the commerciality of their work, how it affects their work, their own artistic ideas and development. Aside from us sitting around as venues and programmers we need to consider who we're programming as well in this relationship." (Organisation)

Quote 131: "... there may be artists who say I don't want you [receiving venue] to share my data because I need to sell that show – if you've got really bad results, I need to sell the show and need to work on it ... I think that's a conversation that wasn't really had. When we signed up to it, we realised there was another level, of us then having to go to artists to say we'd like to do this on your show, are you happy for us to ask these questions, and then are you happy to share that data ..." (Organisation)

Data privacy, data ownership and the Freedom of Information Act

There are a number of different options, some working together and some by themselves; there are cost implications and data ownership implications, resulting in the need for a discussion of the most effective means of collecting the data without a burden to the cost.

Concern was evident from across the consultation about who had control of the data, which ultimately rests upon who owns it and the agreement the owners have with the NPOs/MPMs that share their data. Given the potential commercial sensitivities⁴⁷ that release of particular data could result in and the potential gain that owners of aggregated data could leverage, if the QM initiative is to continue to proceed, absolute clarity will need to be established about ownership, about permissions for release of data, about the degree of anonymity that needs to be assured, and about protocols for the release of information. Participating organisations had been provided with clear guidelines around data ownership and use for the national trial phase but concerns remained. Consideration also needs to be given to whether the means of collecting and hosting the data should be proprietary or platform agnostic, open-source or a form of 'freemium' model, or whether the data can be collected by organisations by whichever means they chose and submit to the Arts Council in a standardised manner for dissemination.

We also note that there is a further concern that has implications for participating organisations and for the Arts Council itself, in that if a public body holds information, there is the potential for release of that information under the Freedom of Information (FoI) Act except where an exemption applies. 'Datasets' and 'statistics' are also subject to specific conditions. Consequently, it is necessary that the Arts Council ensures that participating organisations' concerns about data ownership and usage and the Arts Council's data analysis objectives can be reconciled within the context of the statutory framework of the Act.

Measured enthusiasm for adoption is contingent upon further refinement

⁴⁷ This is also discussed further in Section 4.4.



The consultation has identified an appetite amongst participating NPOs/MPMs for finding a solution to assess quality and in developing the QM framework so that it is more effective. If practical and process challenges are resolved and a clearer intention of purpose articulated, greater enthusiasm for adoption will be engendered. The most commonly reported deterrents for adoption (and conversely encouragements for adoption/ways to address deterrents) reflect those raised through both quantitative and qualitative data and are summarised in the table below.

Deterrents for adoption	Encouragements for adoption or addressing deterrents
<p>Value and benefits are unknown due to a lack of understanding of how to use the data and the uncertainty about the overall purpose for both organisations and the funder</p>	<p>Organisations need further convincing of the cost (including time) versus the benefits of participation – including a clearer indication of how to use the data and a clearer articulation of how the Arts Council will use the data/what the data should be used for.</p>
<p>The current survey design functionality and accessibility of the CC platform, both front-end and back-end</p>	<p>Organisations expect further development and refinement of the platform before it is deemed fit for purpose e.g. clearer reporting mechanisms, higher levels of sophistication with survey design and increased inclusivity through addressing front-end accessibility in particular</p>
<p>If there was a charge for using the framework/if they had to pay without seeing its value/if they had to pay for it as well as Audience Finder</p>	<p>Organisations are price sensitive and currently the majority are unwilling to pay to use the framework or to would only pay up to £100 per year. Organisations need to see the value before committing to any costs.</p>
<p>Desire to add context or provide explanations and/or justifications to the scores</p>	<p>Organisations and peers express a desire to add narrative into their scores and the resulting data dashboard, and the assurance that the data will not be used/viewed in isolation but as part of a whole, broader story provided by multiple data streams such as the contextual information.</p>
<p>Lack of collaboration between receiving and presenting venues and the administrative burden that touring organisations encounter</p>	<p>Organisations report that greater buy-in from both partners is needed to ensure that assessing touring programmes is less admin-heavy on one partner and can therefore be sustainable in the long term</p>
<p>Lack of confidence in the reliability and validity of the data</p>	<p>Organisations want the self-assessment and peer assessment process reviewed and refined to ensure potential bias is eradicated. Addressing the challenges regarding semantics of metrics to ensure each dimension is clear and interpreted in the same manner would also encourage consultees to more readily adopt the framework. Organisations need convincing that the sampling frame is robust</p>
<p>The existing set of QM dimensions</p>	<p>Organisations want the flexibility to select metrics which are most relevant for their discipline and artistic aims (though they recognise that this may diminish benchmarking capabilities) and an agreement that whilst some metrics could be compulsory, there is not a 'one size fits all' approach</p>
<p>Confidentiality concerns over data ownership</p>	<p>Review the legal applicability of releasing statistical data within the statutory framework of the FoI Act and upon this basis clearly articulate to the sector the guidelines that would apply to release of data, extent of anonymity and data ownership</p>



Deploying the QM framework at the same time as Audience Finder

Organisations want to see some alignment with existing compulsory data capture systems like Audience Finder to reduce capacity burden but an overall solution will need to be found to counteract survey fatigue through responsible sampling

If it replaced the existing AQA system

AQA peers would want to use the QM survey in combination with the AQA existing report format or provide more opportunity within the front-end system for assessors to add narrative

Limited capacity for delivery (especially for non-ticketed or those with combined arts programmes with many 'one off' events)

Organisations want to see greater financial investment provided to support those with limited capacity who have to buy-in fieldworkers and kit, and greater flexibility with, and clarity of, sampling expectations

Not a quick win

Overall it is evident that organisations believe that suitably assessing quality across the portfolio is not going to be a quick win, and certainly not resolved through one national test phase (albeit building on learning from previous iterations). Indeed, consultees (including those who had taken part in previous trials) reported that the QM framework needed further work and investment to refine and develop the individual dimensions and address challenges presented by the framework itself, such as peer recruitment, before the scheme could be successfully rolled out.

Quote 132: “Probably in all honesty you’d need to have 10-15 years of this experience behind you actually can’t start to say there’s elements of this that are quite robust.” (Organisation)

Quote 133: “Give it time...it will work.” (Organisation)

Summary

Overall the sector supports improving quality assessment, including a specific interest in the QM framework, which through the national test phase has already stimulated intra-organisation debate about defining/measuring quality and the consequent use of data. Alongside this good will, there is nonetheless a recognition of the innate difficulty in defining quality and in achieving cross-discipline applicability of a single set of metrics and perhaps more importantly, considerable confusion, anxiety and lack of understanding about what is ultimately intended, despite the Arts Council and CC messaging. This is further complicated by QM leads from participating organisations in this national test phase usually coming from two different departments with differing frames of reference. For example, whilst an artistic director may tend to rely more on perceptions of artistic excellence, a marketing director may be more accustomed to using data to evidence their decision-making. As it currently stands, a number of deterrents and potential emollients have been identified, and any enthusiasm for adoption of QM is contingent upon the range of concerns raised by this evaluation being addressed.



There are significant concerns about data ownership, data use and data anonymity, which the Arts Council needs to resolve and then relay the solution to the portfolio. This needs to be considered with reference to whether the statistical data would fall within the binding statutory framework of the Freedom of Information Act.

For the QM framework to have a chance of demonstrating what could be achieved by this shift in quality assessment and data culture, clearer and stronger articulation of what is intended needs to be relayed to the sector. This needs to be addressed to both programming and marketing functions, so as to engender true buy-in, which will be critical to the programme's success, as it will ultimately be the organisations, working together with the Arts Council and any future service provider that will find solutions to making the programme viable and relevant.

Lastly, for the programme to be successful, patience will be required as the sector faces a period of culture change towards using data and assessing quality in a focused, rigorous and more comparable way. Such a quantum change needs to occur alongside development of flexible process solutions and sector skills development, so that the value of QM can be better realised and not undermined. There are potential lessons to be learned from the Australian experience that suggest that this change can be wrought. In short, this will not be a quick win but one that will require perseverance and good will.

Case Study: Learning from the experiences of Western Australia

In 2010, the Department of Culture and the Arts (DCA) in Western Australia identified a lack of data analysis in the sector, and in witnessing opportunity costs on both a micro (organisational) level and a macro (sector) level, set out to develop quality metrics in partnership with the sector.

DCA's aims were fourfold:

1. To equip artists and arts organisations with more information, in a cost effective and contemporary way, in order to help them self-assess against their own goals and to provide artists and organisations with a means of evidence-based decision making (both artistically and administratively).
2. To stimulate discussion around quality of artistic and cultural activities, and to provide a platform for debate.
3. To address the opaqueness of public funding decisions and to find ways of engaging the public in (or democratising) the assessment of quality through a standardised approach and common language.
4. As part of their public value measurement framework, DCA sought to give a full, evidence based account of value through utilising new technology to generate meaningful big data.

DCA approached implementation head-on.

5. In seeking to find solutions to these aims, DCA recognised that the solution through the metrics would need to come from the sector itself, and put a call out to tender for a standardised technical solution.
6. The rationale for going to tender was because no system existed and the cost of collecting this data and analysing it by other means was prohibitively expensive and onerous on both the sector and DCA.
7. DCA has been investing in developing and trialling the metrics directly with the sector, which is still ongoing and developing.

There are a number of common themes and issues shared between the experiences from England and Western Australia, at times offering potential solutions.

8. Organisations were initially sceptical but for the most part DCA noted that by and large they now see value in it, citing that " 100% of our organisations who used it once are now are voluntarily using it".
9. DCA recognised that organisations would initially have a difficult time in meeting the financial and human resources required to participate in QM. To address this, DCA covered the financial cost of QM and provided staff support for the initial implementation of the tool, the system and its web portal, to help organisations along the way. After the

introductory training, the organisations were then in charge of their Culture Counts evaluations. Organisations report that the system is self-explanatory and user friendly and welcome the ease of creating new evaluations once having set up their ideal combination of surveys, peer assessment groups and questions.

10. Prior to QM, funded organisations were required to report on quality in entirely un-standardised ways (such as by providing long narratives, artistic statements and press clippings which were unable to be aggregated to give an assessment of value). Today, however, many organisations report having greater confidence in their reporting on quality as a result of QM replacing the previous model, and the Culture Counts platform allows automatic reporting on quality and provides raw data to be combined with other survey tools. Users and DCA were able to access their quality data automatically on their programmes combining artist, peer and audience voices.
11. For any privacy concerns, DCA adheres to strict data protection measures and communicates clearly that no individual peer or audience assessments will ever be visible by anyone, and that all data is anonymised and presented in aggregate only.
12. In the early days, organisations felt that the data was less reliable and thus less useful, however, over time as they amassed more data, the data became increasingly reliable and thus useful in their decision-making.
13. To mitigate early concerns over the misuse of data, DCA clearly communicated the aims and intentions of QM: First and foremost, QM is to benefit organisations in helping them assess against their *own* goals and assumptions; secondly, QM will not be used compare organisations on their absolute metrics (and any attempt at doing so would be erroneous); thirdly, it's about creating a discourse where the debate is as important as the outcome. Whilst all data remains anonymised and in aggregate only, there may be an opportunity in the future to find further public value in the data through the government's open data policy and the academic community (remaining confidential).
14. Organisations sought flexibility in the QM dimensions, and to address this, the pre-event self-assessment is intended for organisations to set their own assumptions and to enable them to narrow in on the key metrics in context. In providing greater flexibility, a minimum core set of metrics were required from organisations and the remaining were made optional. Based on their experiences, however, Colin Walker of DCA noted that if they were to do it all over again they would have made all of the QM dimensions mandatory for all organisations "because of the 'what if' question... and because it's about the debate around the metrics" to challenge thinking and stimulate discourse. Applauding the decision in England to make all of the dimensions mandatory, Walker noted that "the cleanliness of everyone using the same metrics on the same scale will be an advantage in terms of big data in the long term".
15. Organisations already struggling with capacity issues to collect different types of data for different purposes or systems (e.g. box office data, demographic data and quality metrics data) have demonstrated a demand for an integrated approach. Recognising not only this

demand but also the opportunity for greater use of big data, DCA anticipates a market-driven approach to integration across the various data systems. Walker notes that at the core of quality assessment needs to be standardisation on the scale used for the QM without which comparison, benchmarking and aggregation is not possible. Of all the data collected, Walker notes that it is the quality data that the boards typically want to see and discuss, as this is what goes to the heart of the organisations, and supersedes the desire for integration.

16. Organisations have had a deficit in some of the skills needed to fully realise the potential benefits of the data. Over time, as the sector becomes increasingly data savvy and organisations become increasingly data driven, the skill sets required are developing amongst the workforce.
17. For organisations with limited access to hardware such as WiFi, tablets and mobile connectivity (e.g. dongles), DCA provided necessary kit. In some instances, organisations continue to use paper for the audience surveys on the same scale. For some events, audience surveys were distributed by email subsequent to the event. DCA has also done analysis on the timing of the survey and have found that assessments are consistent whether they occur immediately after or weeks following the event, reporting that there is no 'halo effect'. Recognising the demand and opportunity for greater online and offline capabilities, market-driven approaches are in development by providers.



5. Closing remarks

The national phase trial of the Quality Metrics framework has seen 150 portfolio organisations take on the considerable task of testing the framework (on top of the other initiatives in which they may be involved). Great efforts have been undertaken by the sector, by the Arts Council and by the platform providers in rolling out this programme. This evaluation highlights the desire of the sector to have a framework that works. It also provides the Arts Council with an assessment of participating organisations and peer assessors of their experience of the trial and of their insights from a wide set of perspectives.

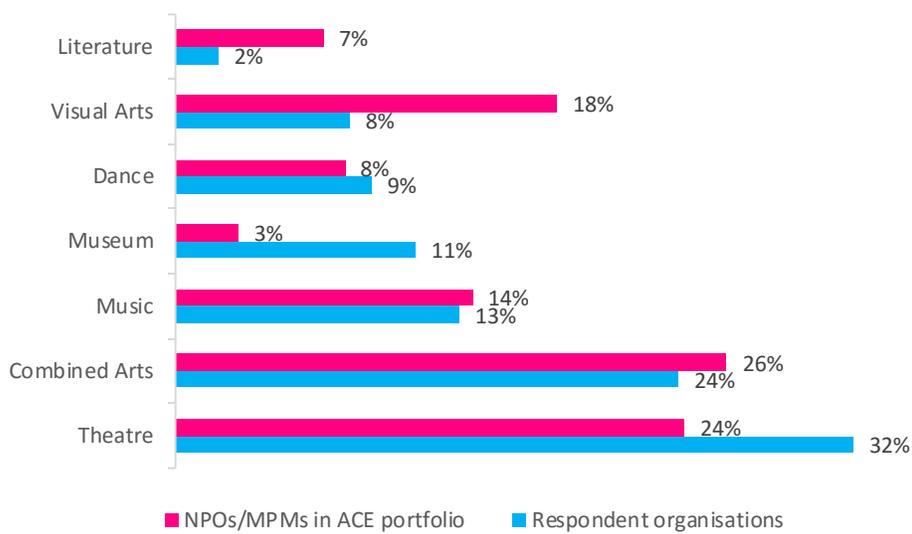
Much has been learnt by the sector and it has stimulated conversations within organisations about assessing quality in a reproducible way and about the intersection of quality and data. The Arts Council will have learnt much from the sector about what is effective and what is not, and should have a sense of the challenges ahead rendering the framework viable. The solutions and compromises necessary will in large part come from the sector as they lead the project forward supported by the Arts Council in making the framework effective for a broad range of organisations and disciplines. As the experience of Western Australia shows, public investors working with sector can bring about an evolution of the framework that better satisfies the sector and public investor.

An evidence driven, decision making future is there to be grasped – particularly if concerns about QM are addressed and capacity is developed within the sector to couple data competence with the already strong commitment to cultural excellence.

Appendix A: Participating Organisations' Survey Report

Profile

Figure 19: Respondent organisations' primary discipline



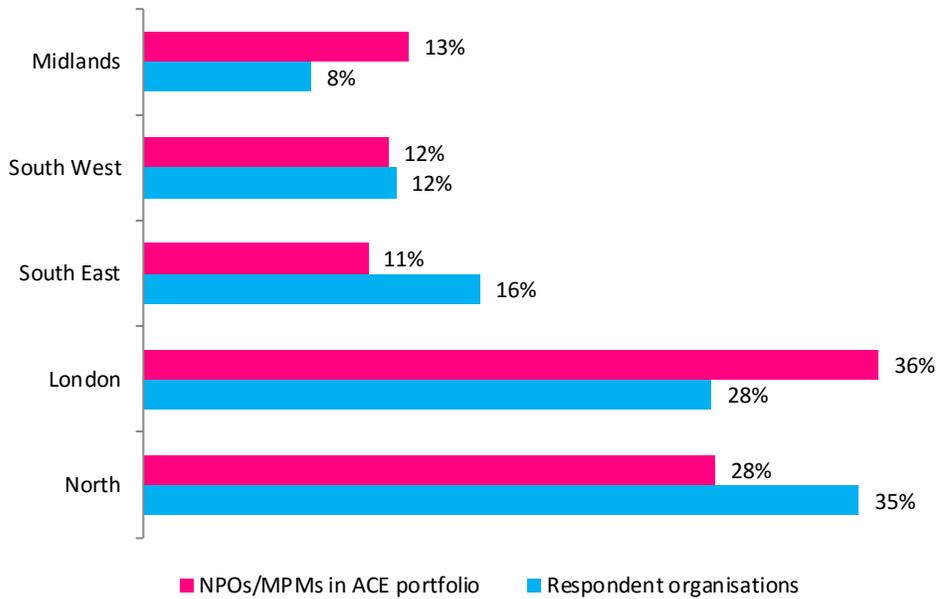
Count = 97

Figure 20: Respondent organisations' size by number of employees



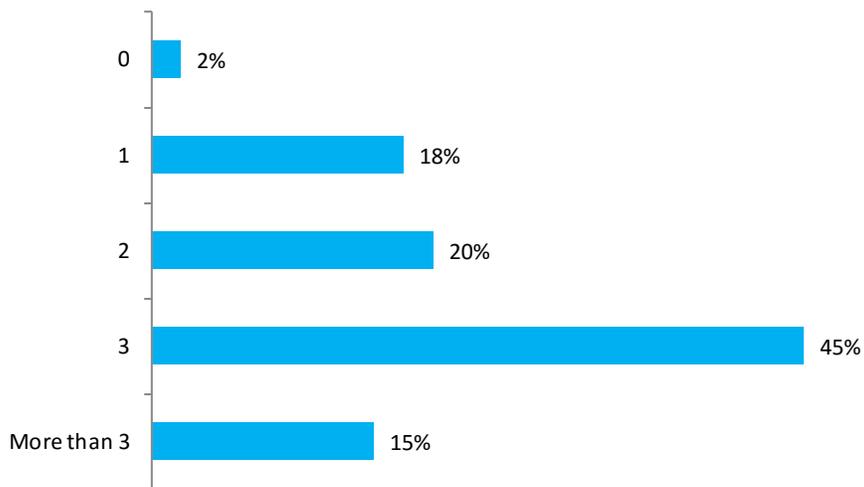
Count = 97

Figure 21: Geographic distribution of respondent organisations



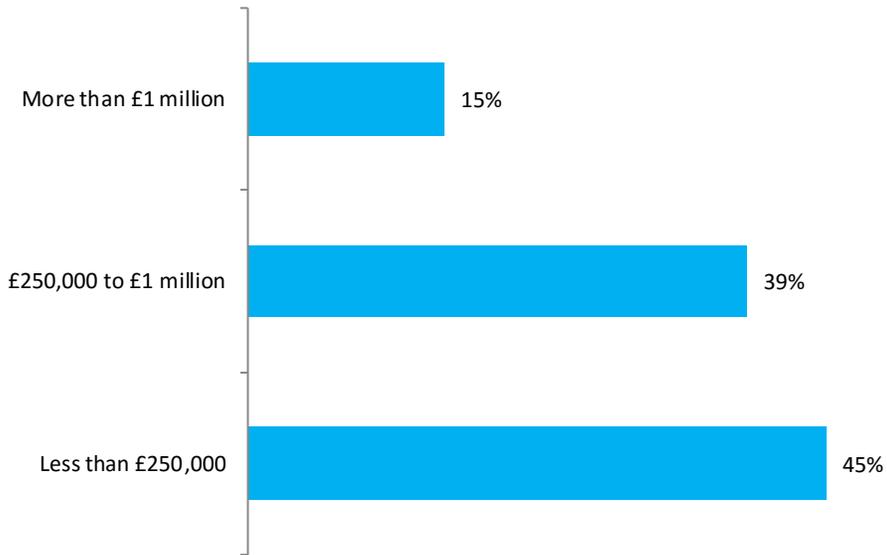
Count = 97

Figure 22: Number of events assessed by respondent organisations for the national test phase



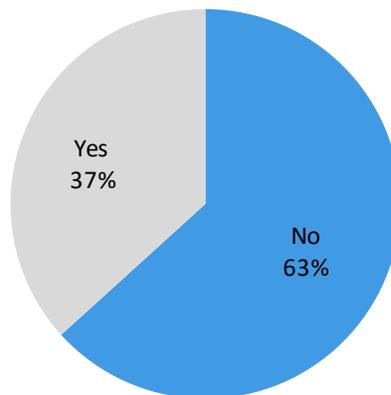
Count = 97

Figure 23: Level of portfolio investment in respondent organisations per year



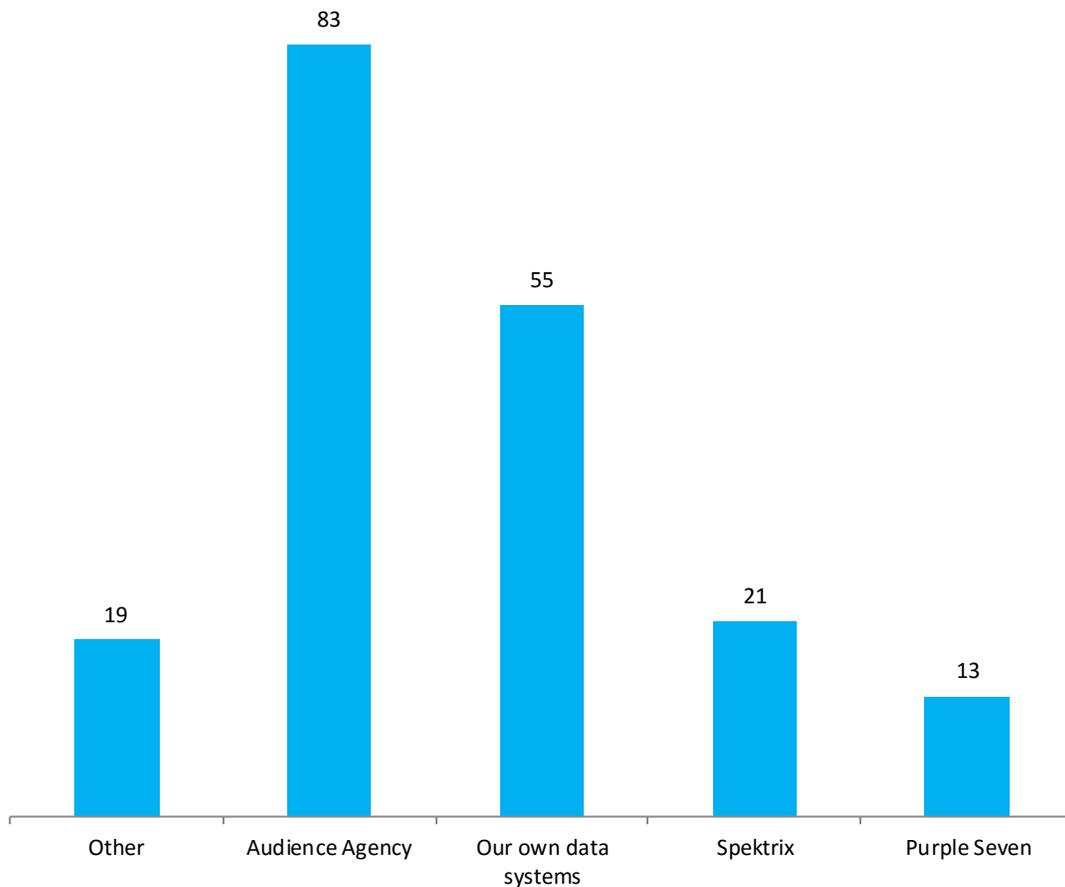
Count = 97

Figure 24: How many organisation survey respondents were personally also peer assessors



Count = 98

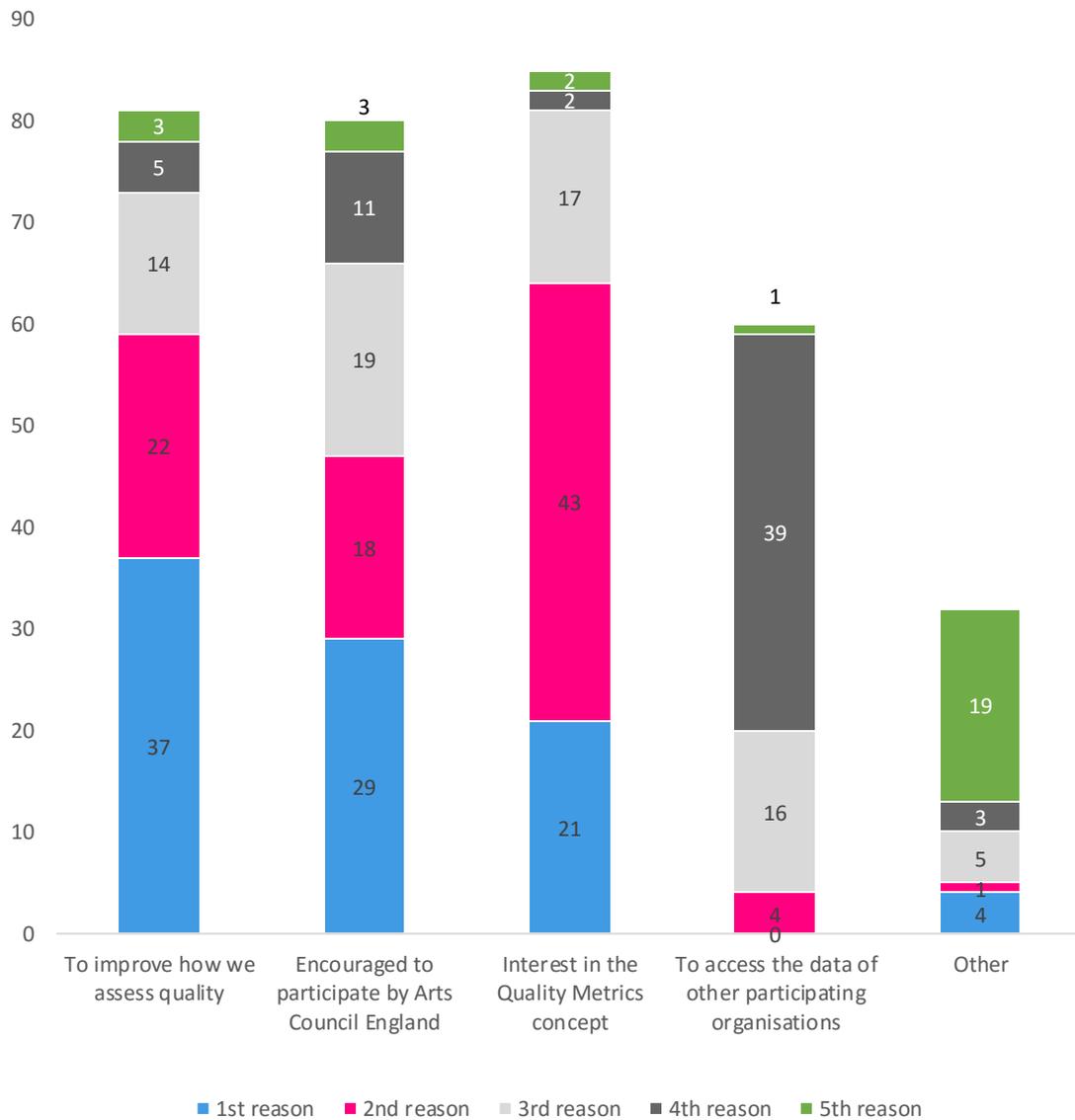
Figure 25: Other data collections tools being used by respondent organisations (frequency)



The other data collection tools that survey responders indicated that they used were: Advanced Donor Strategy, MailChimp, an interviewing agency, ALVA, 'as a producer we have to work in partnership with data collectors', Audience Experience Survey via Small Venues Network, evaluation surveys, Google Surveys, independent market research, Morris Hargreaves, comments cards, Patronbase, Purple Seven via our venues, Survey Monkey, Tessitura, TRG Arts, and Visitor Verdict.

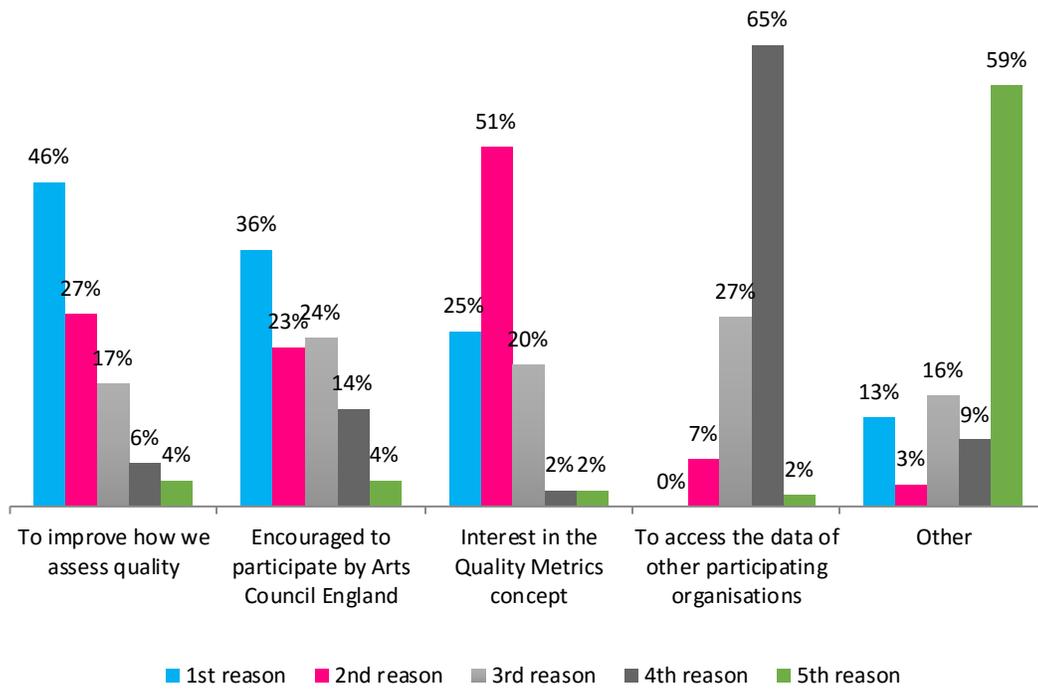
Analysis

Figure 26: Motivation of respondent organisations to participate in national test phase (frequency)



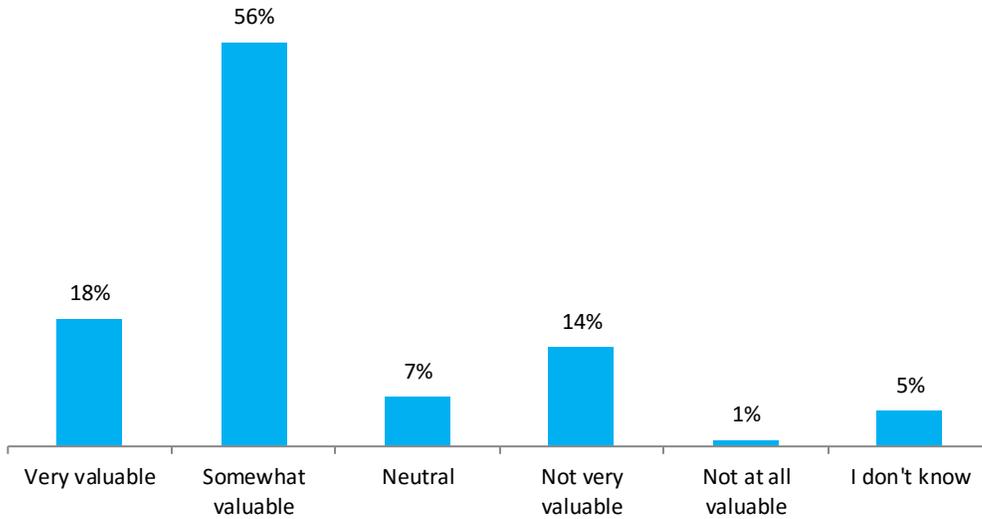
Count = 81, 80, 85, 60, 32

Figure 27: Motivation of respondent organisations to participate in national test phase (percentage)



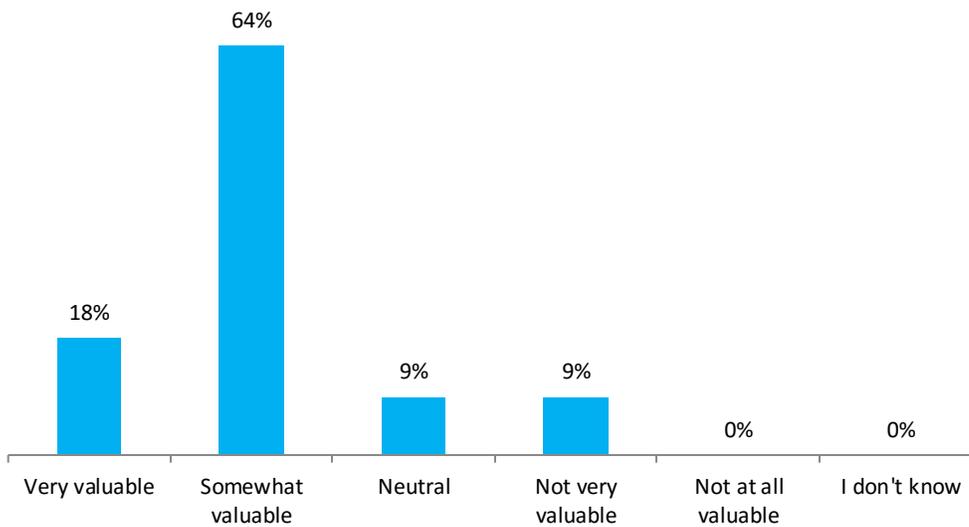
Count = 81, 80, 85, 60, 32

Figure 28: Value of QM concept to respondent organisation



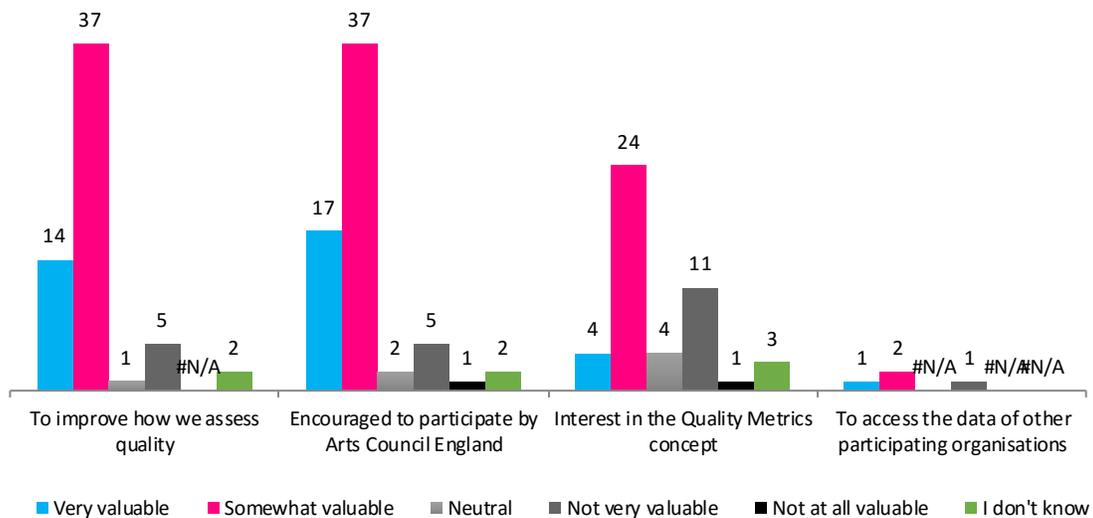
Count = 102

Figure 29: Value of QM concept to respondent organisations with a diversity focus



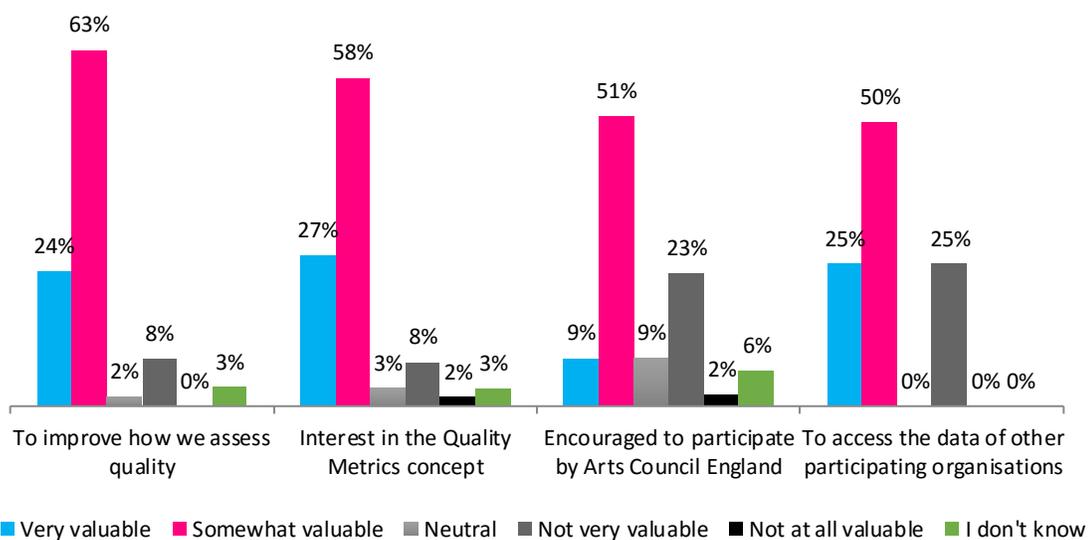
Count = 11

Figure 30: Value of QM concept to respondent organisations when cross-tabulated with motivation (primary or secondary reason for participating in national test phase) (frequency)



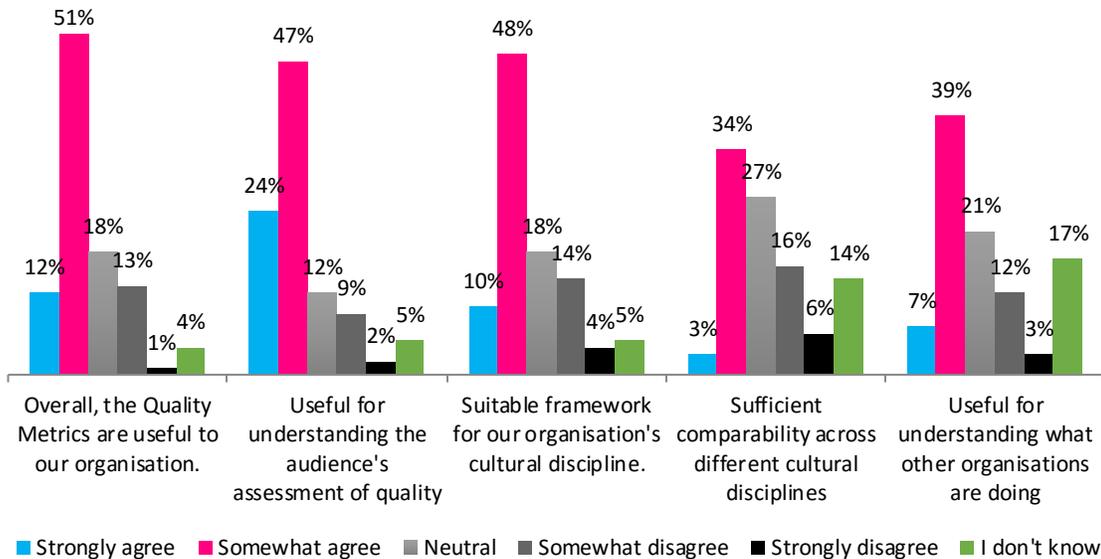
Count = 59, 64, 4, 47

Figure 31: Value of QM concept to respondent organisations when cross-tabulated with motivation (primary or secondary reason for participating in national test phase) (percentage)



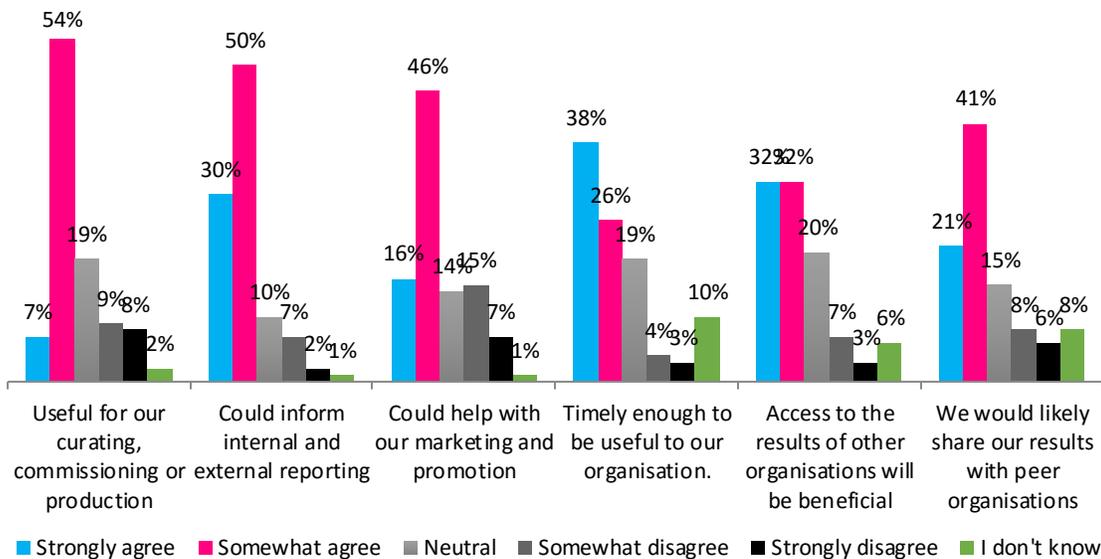
Count = 59, 64, 4, 47

Figure 32: What respondent organisations thought of particular elements of the QM framework



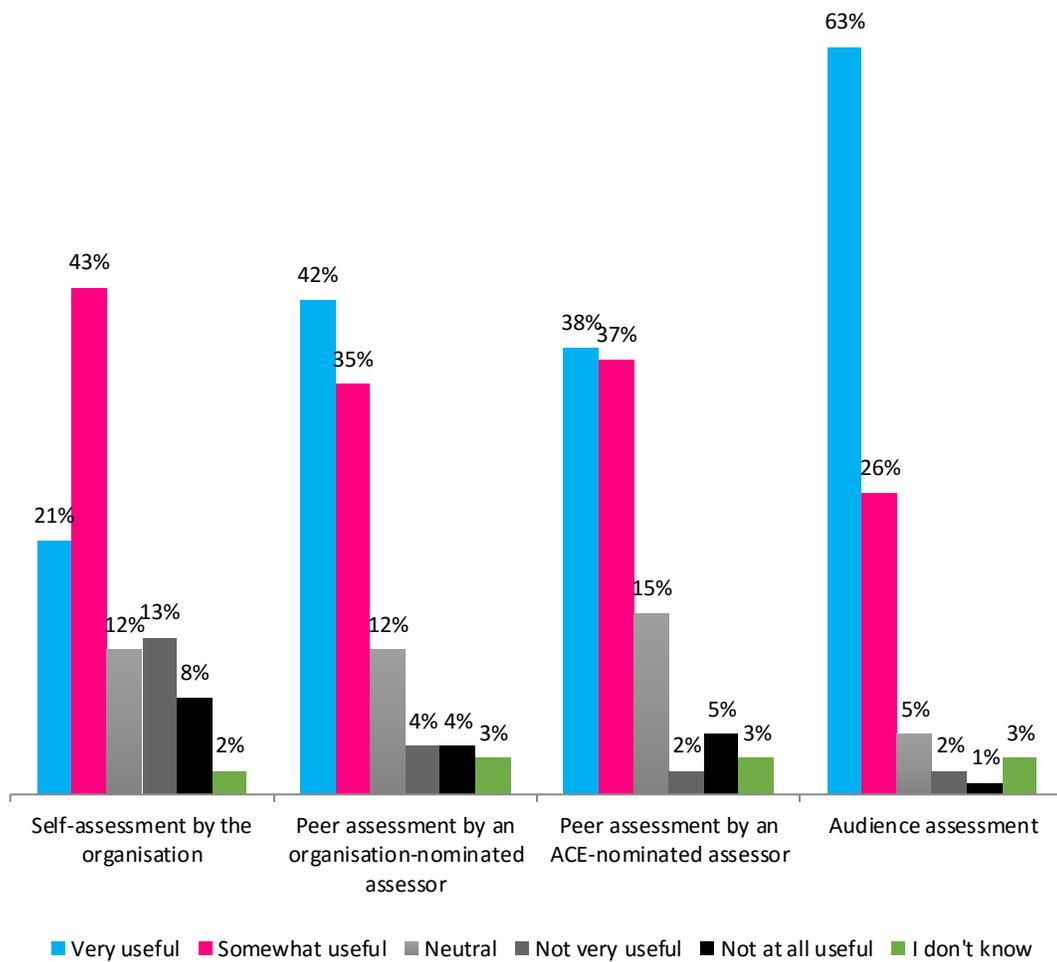
Count = 98

Figure 33: In what way could the QM framework help the respondent organisations



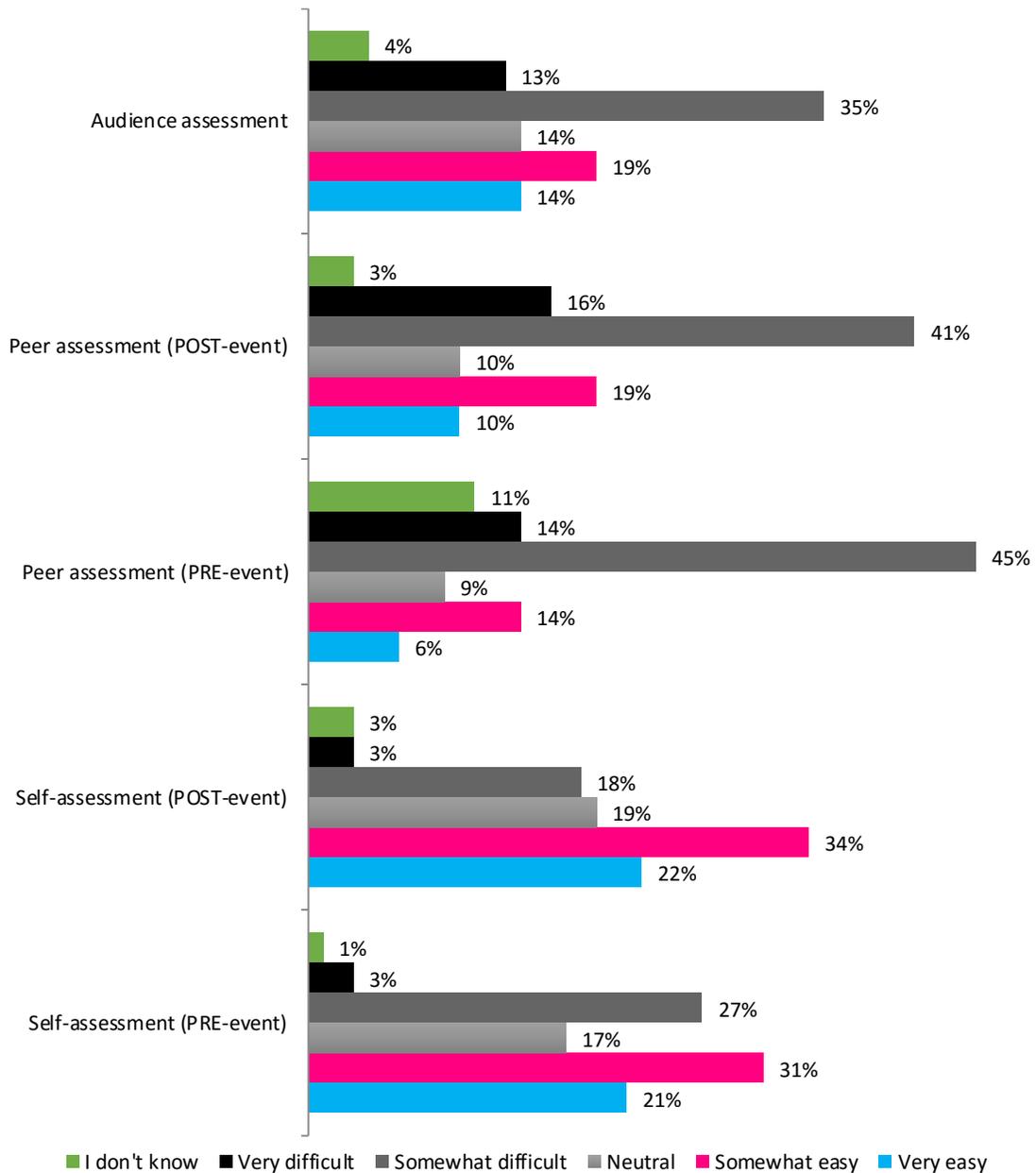
Count = 98

Figure 34: What respondent organisations thought of the triangulation elements



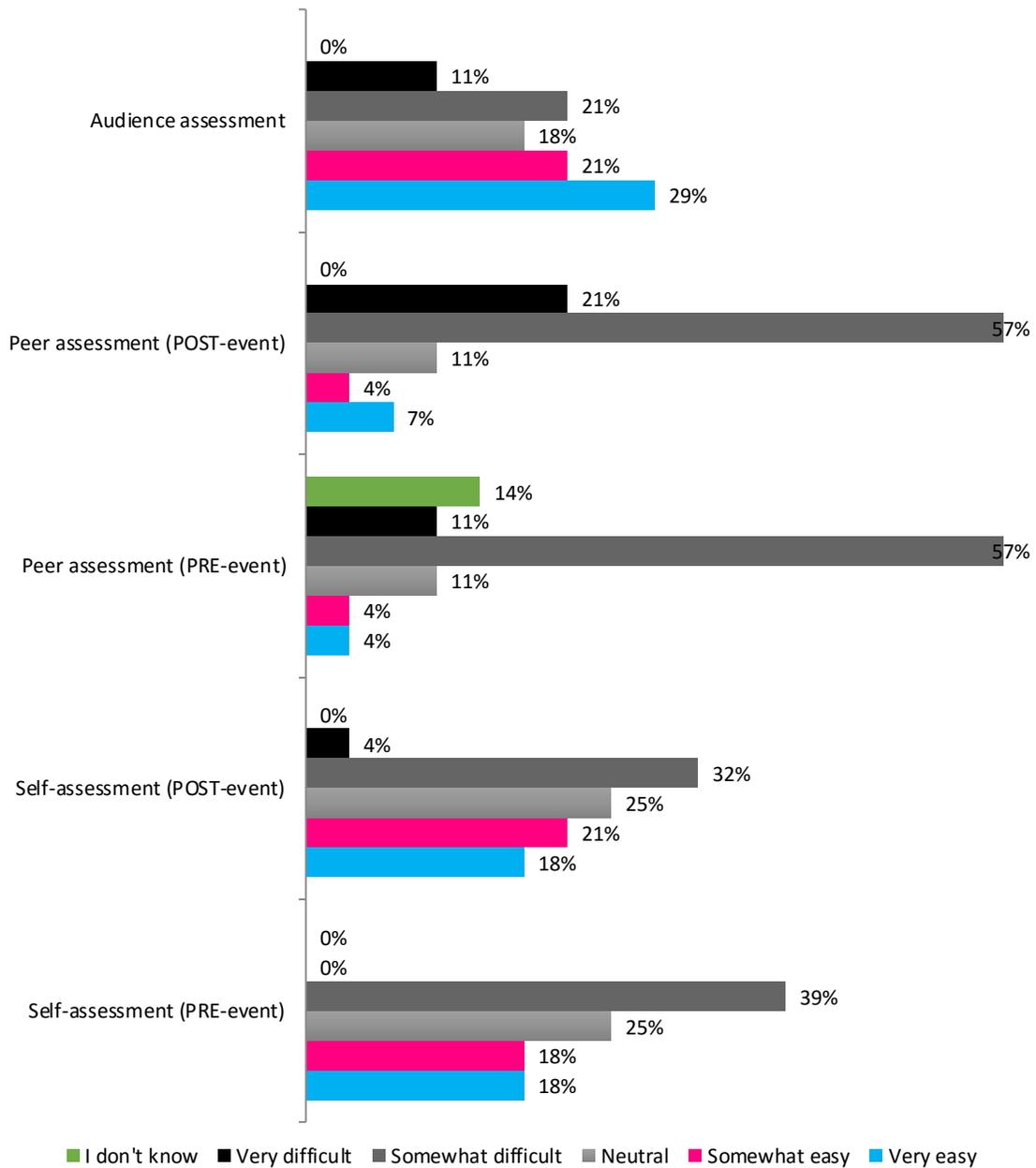
Count = 98

Figure 35: The administrative burden for coordinating and/or conducting each element of the triangulation



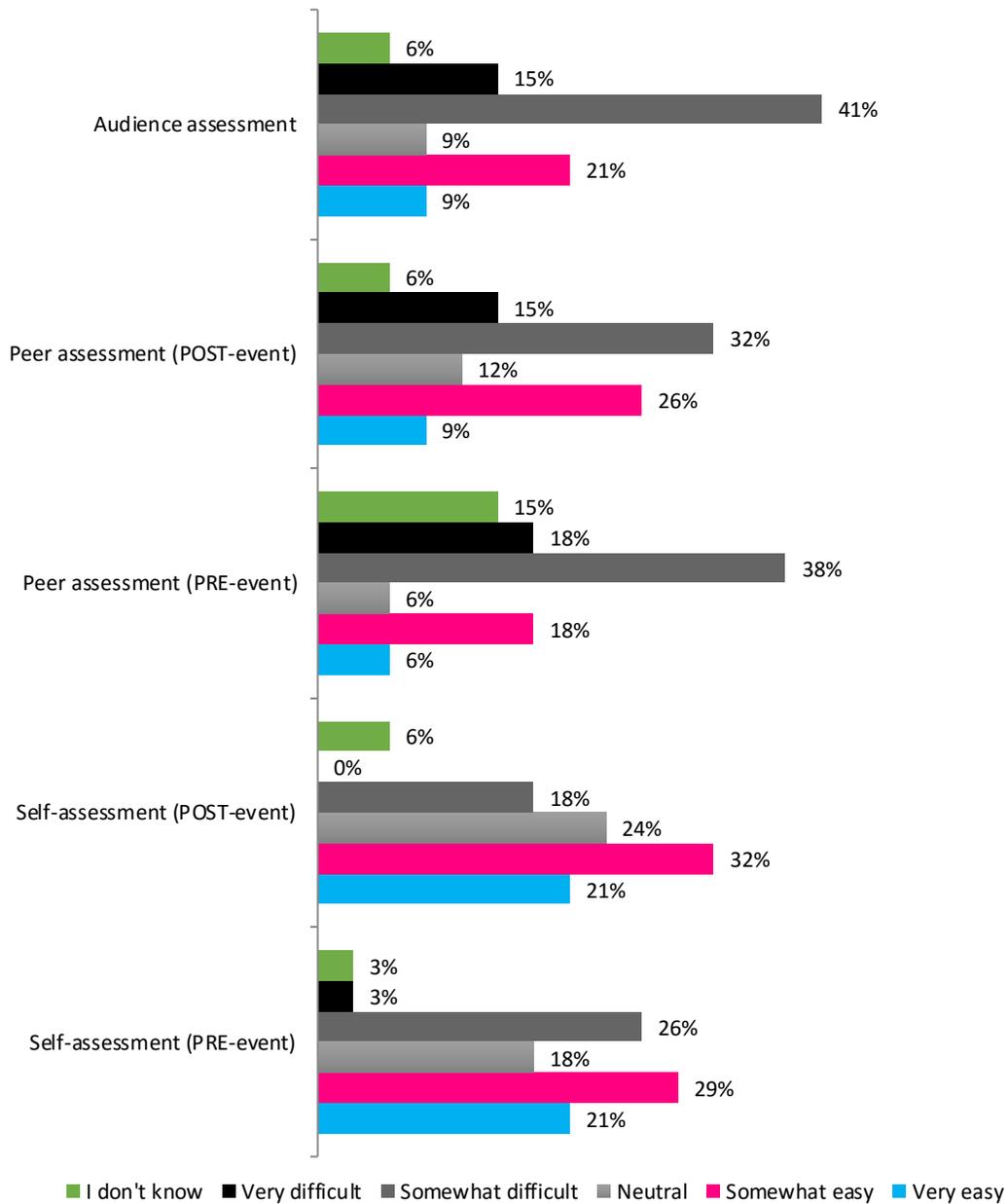
Count = 98

Figure 36: The administrative burden of coordinating and/or conducting each element of the triangulation (Organisations of 50+ permanent staff)



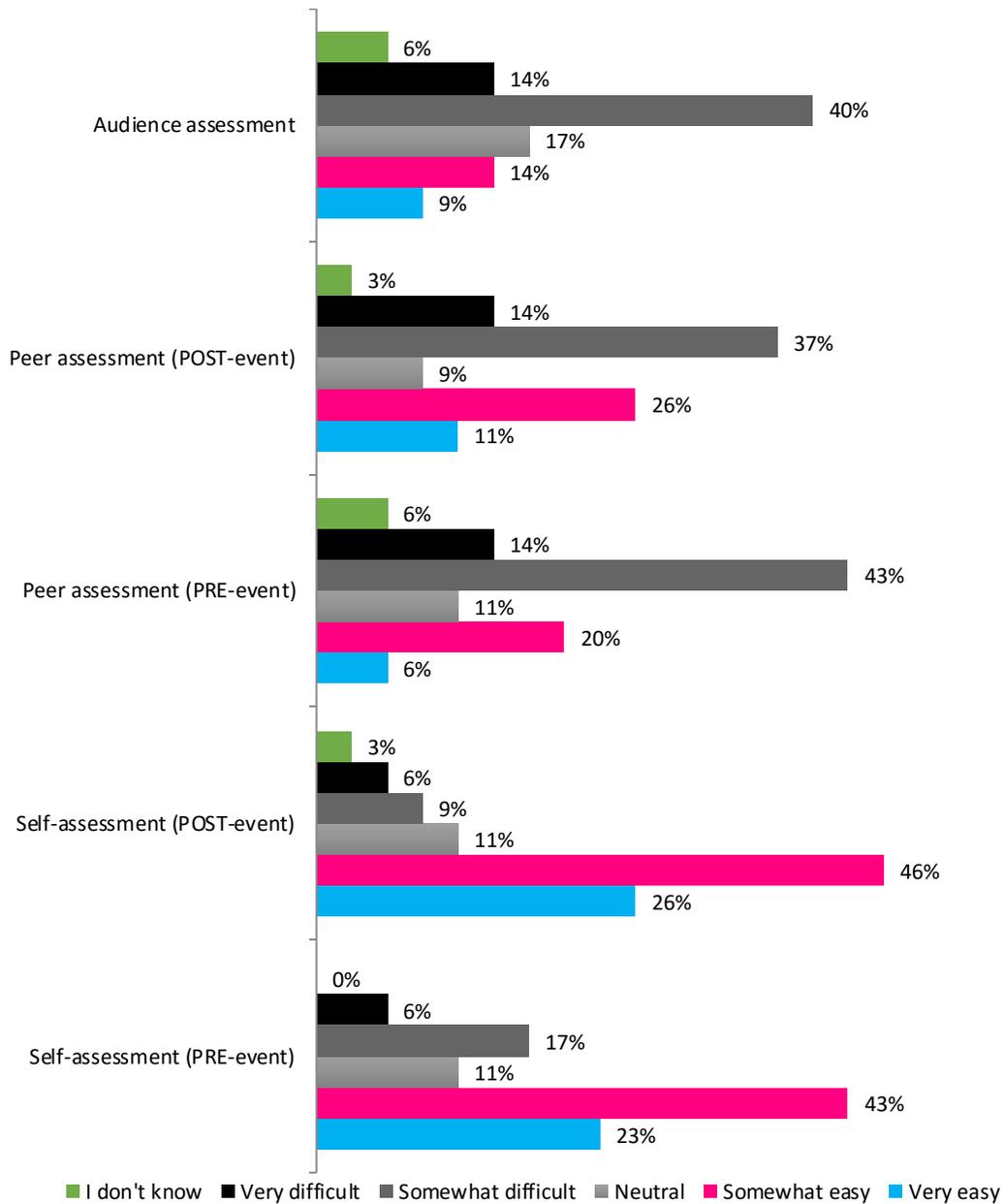
Count = 28

Figure 37: The administrative burden of coordinating and/or conducting each element of the triangulation (Organisations of 11-50 permanent staff)



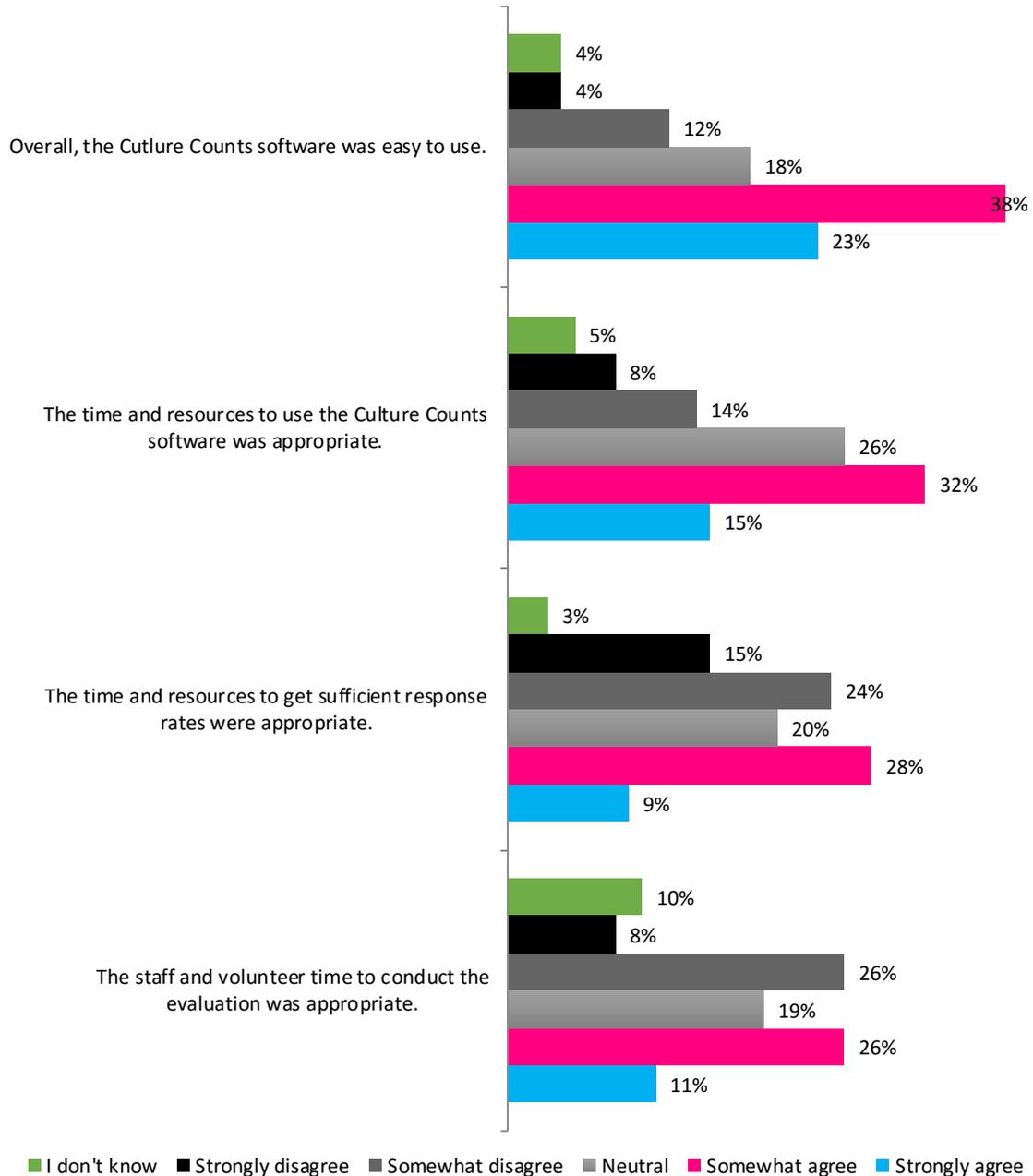
Count = 34

Figure 38: The administrative burden of coordinating and/or conducting each element of the triangulation (Organisations of 1-10 permanent staff)



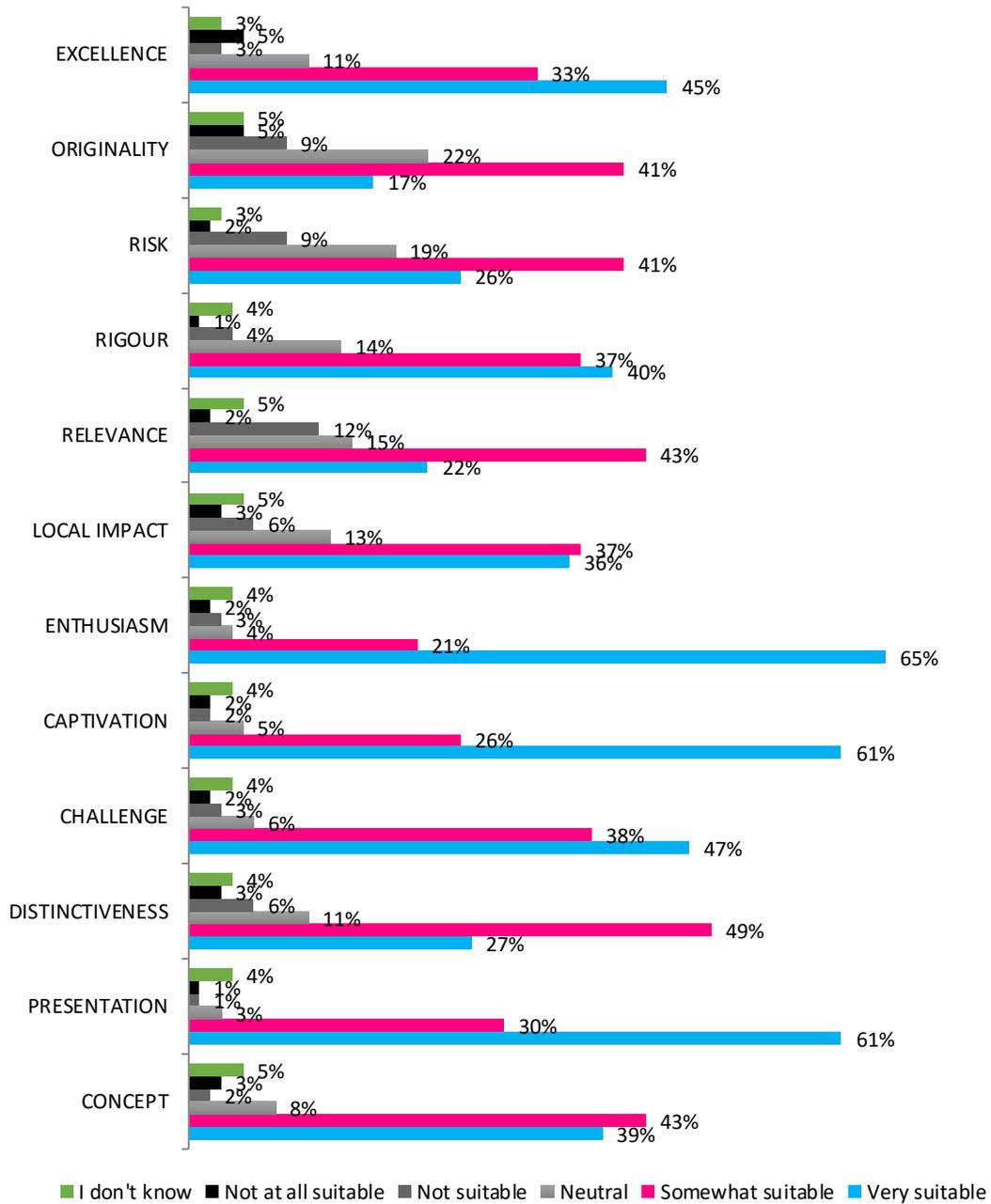
Count = 35

Figure 39: Further insight into what respondent organisations thought of the process



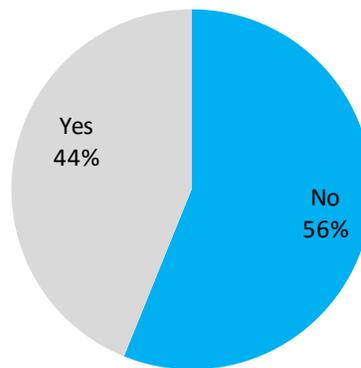
Count = 98

Figure 40: Suitability of individual QM dimensions for measuring quality



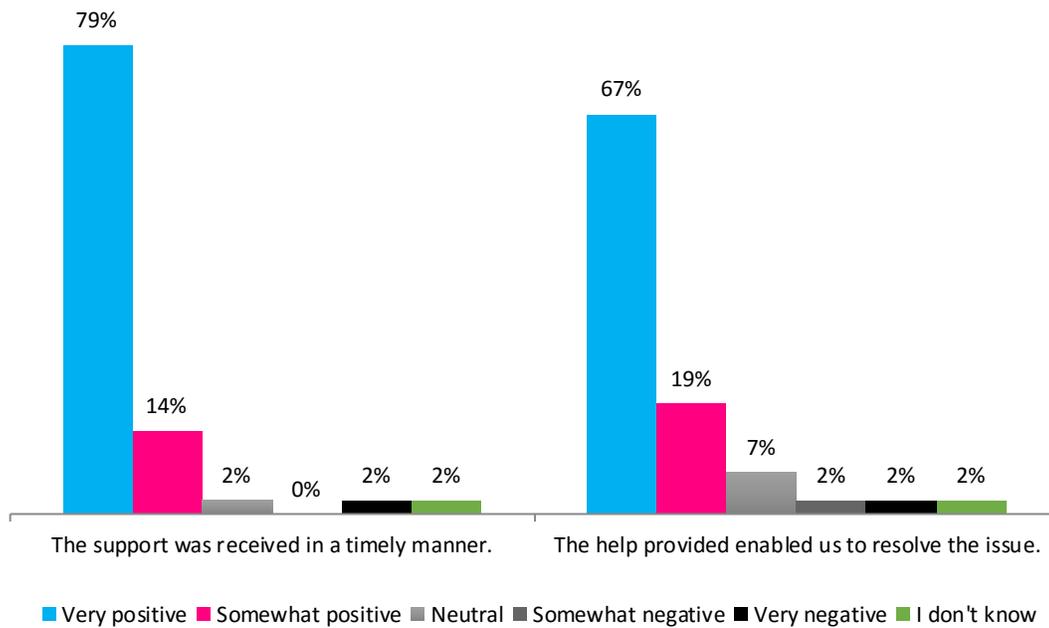
Count = 98

Figure 41: Was help sought from CC for resolving platform problems that organisations had



Count = 98

Figure 42: Timeliness and quality of the help received in resolving problems with the CC platform *

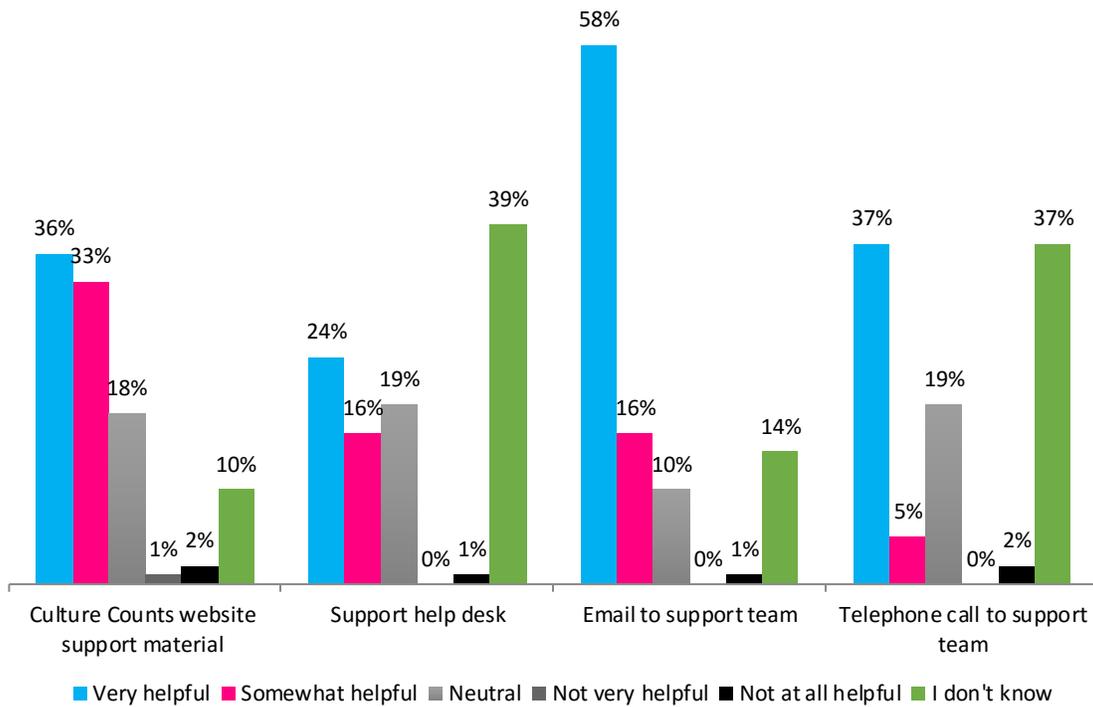


Count = 43

* These are results of the the 44% of respondents that had sought help from CC (see **Figure 41**).

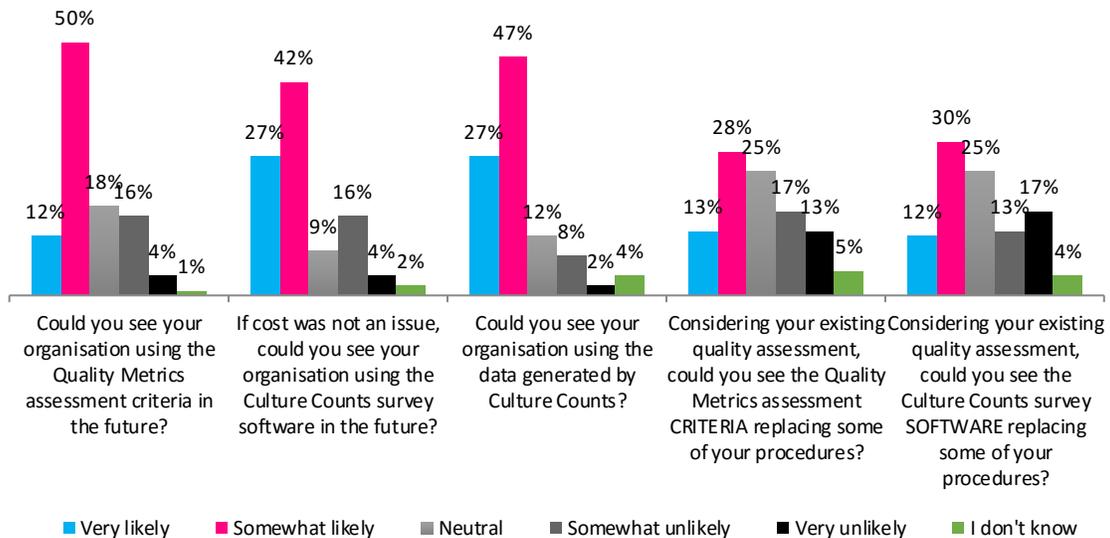


Figure 43: Helpfulness of each medium of support available for CC platform



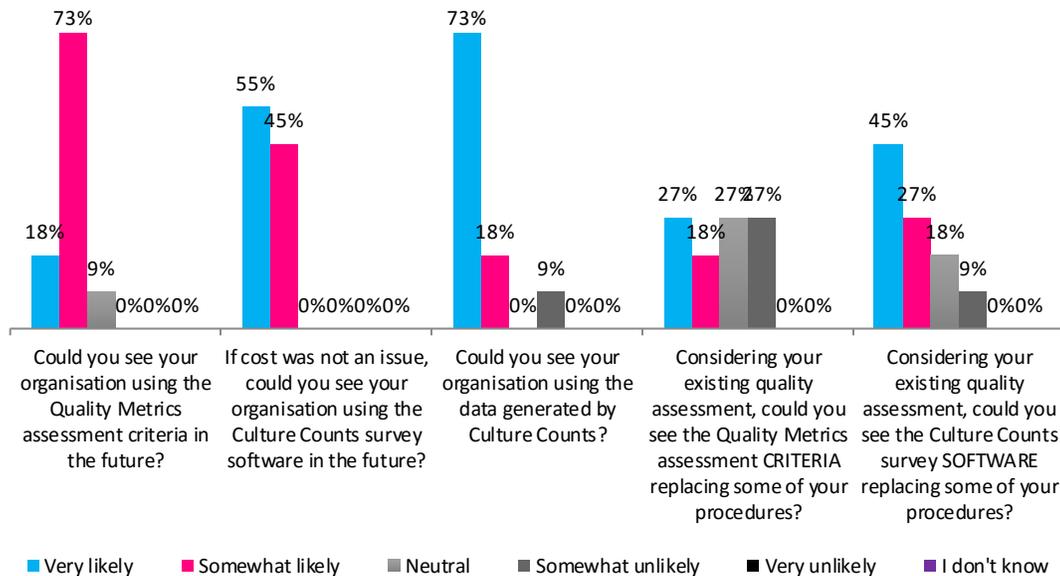
Count = 98

Figure 44: Likely future intentions of respondent organisations



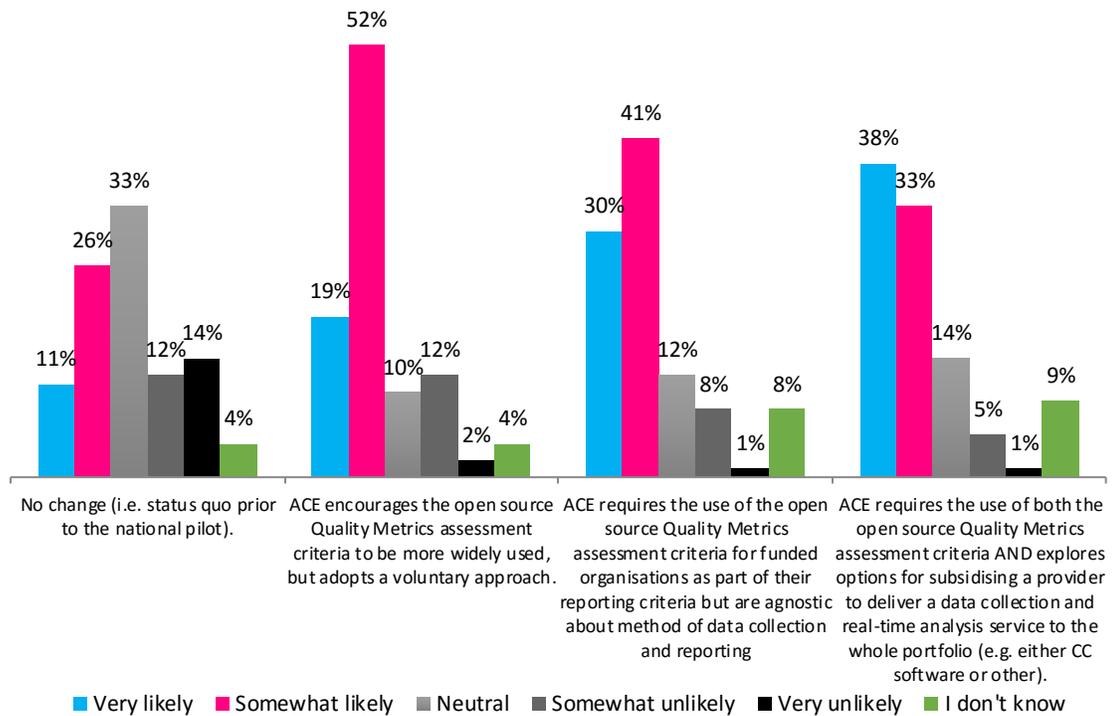
Count = 102

Figure 45: Likely future intentions of respondent organisations with a diversity focus



Count = 11

Figure 46: Respondent organisations appetite for QM under particular future scenarios

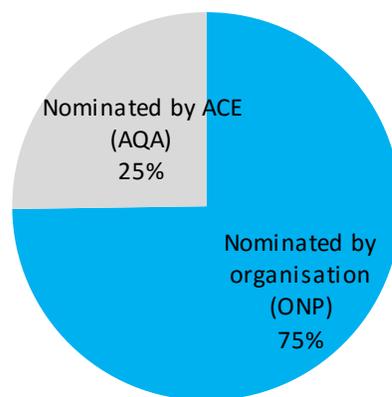


Count = 98

Appendix B: Peer Assessors' Survey Report

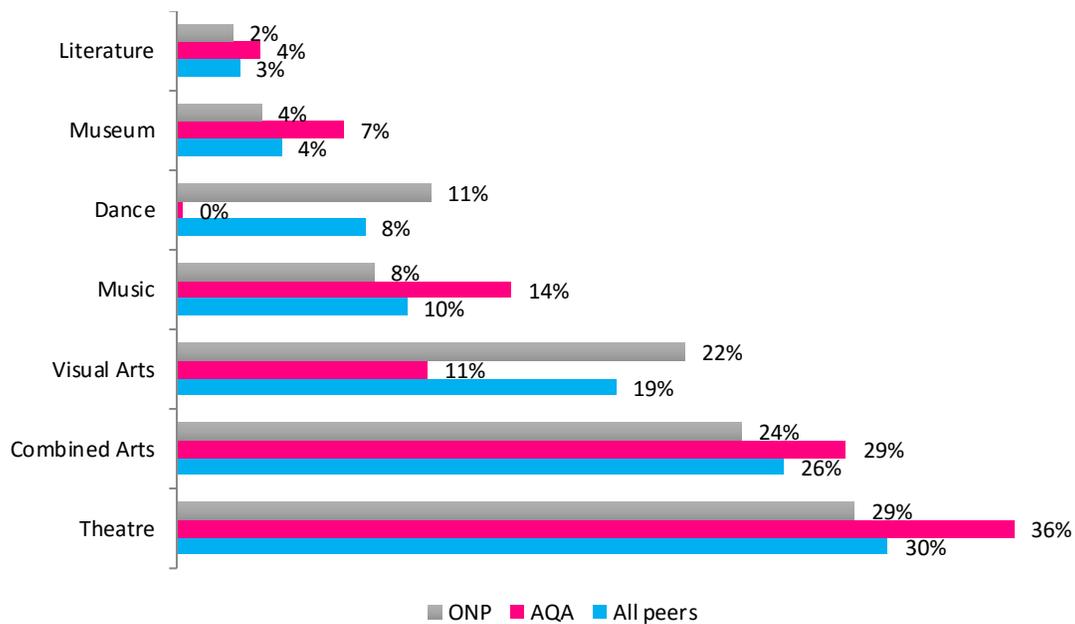
Profile

Figure 47: Type of peer assessor



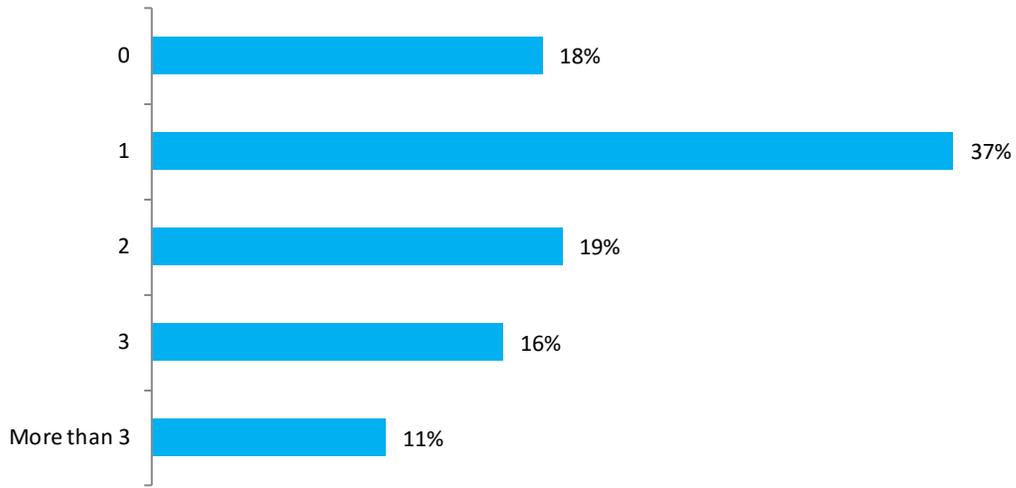
Count = 111

Figure 48: Primary cultural discipline or area of expertise of peer assessors



Count = 83, 28 & 112

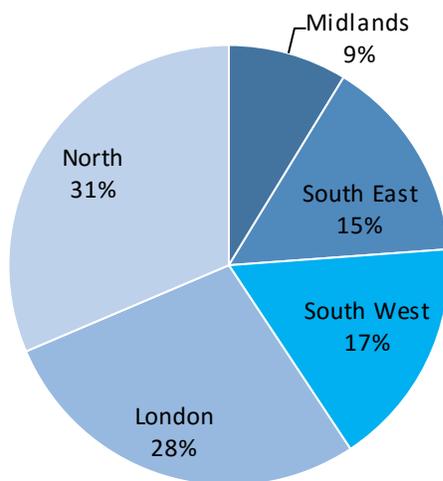
Figure 49: Number of events assessed by all peer assessors



Count = 112

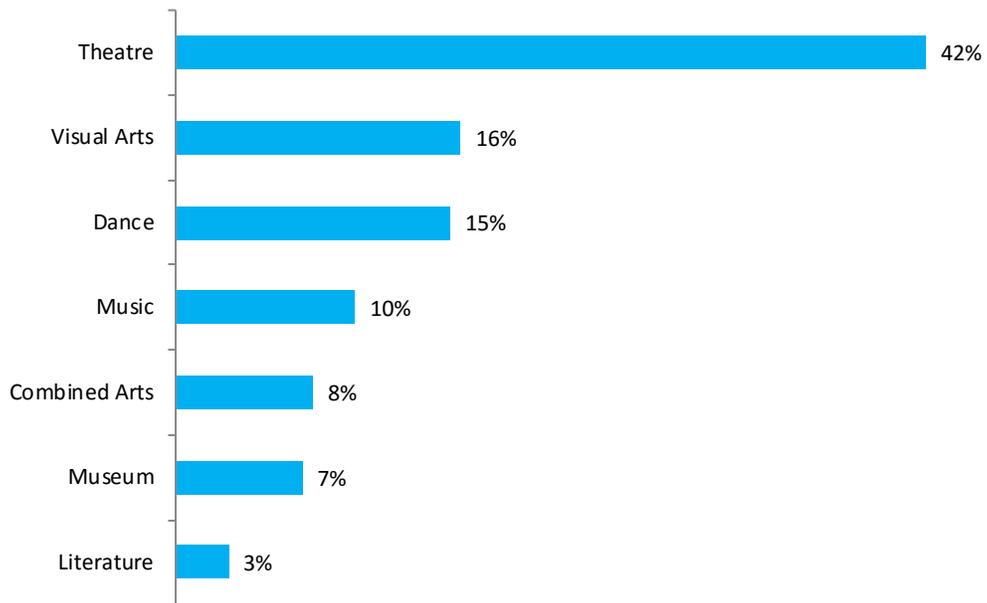
* Approximately one in six peer assessors at the time of the survey had not conducted an assessment, as the survey was conducted partway through the national phase.

Figure 50: Location of events assessed by all peer assessors



Count = 172

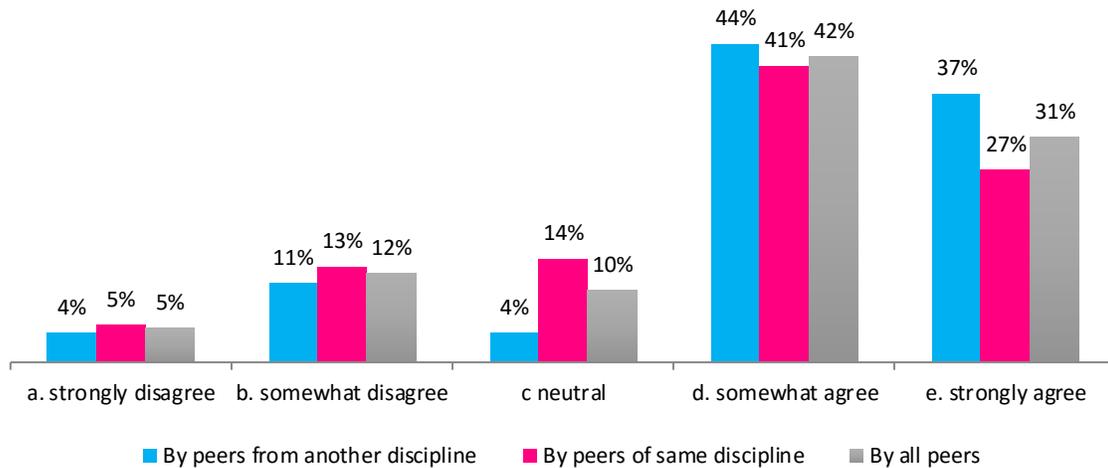
Figure 51: Discipline of events assessed by all peer assessors



Count = 171

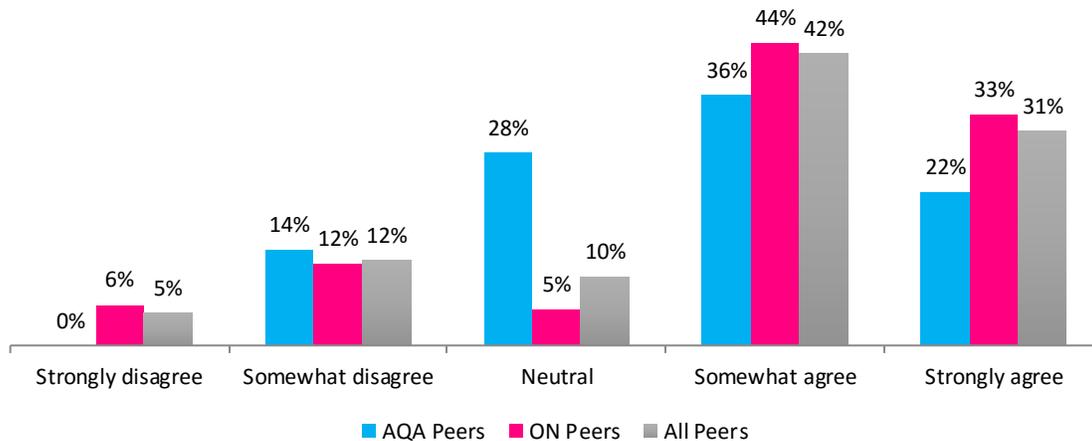
Analysis

Figure 52: Suitability of QM dimensions for assessing the event attended as judged by type of peer (of the same or different discipline)



Count = 73, 98, 171

Figure 53: Suitability of QM dimensions for assessing the event attended as judged by type of peer (whether AQA or ON peer)



Count = 36, 135, 171

Figure 54: Suitability of QM dimensions for assessing particular disciplines as judged by all peers (frequency)

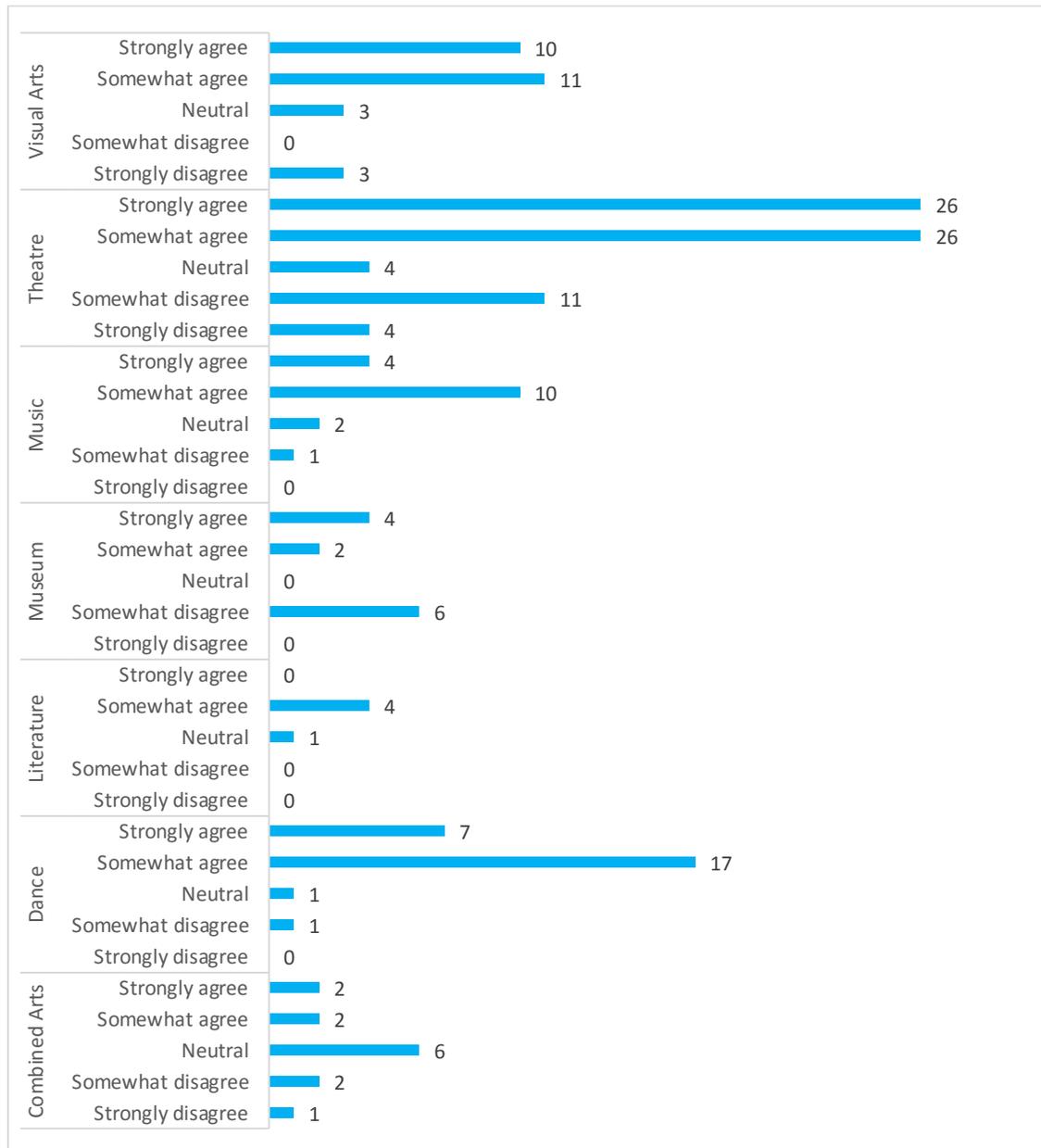
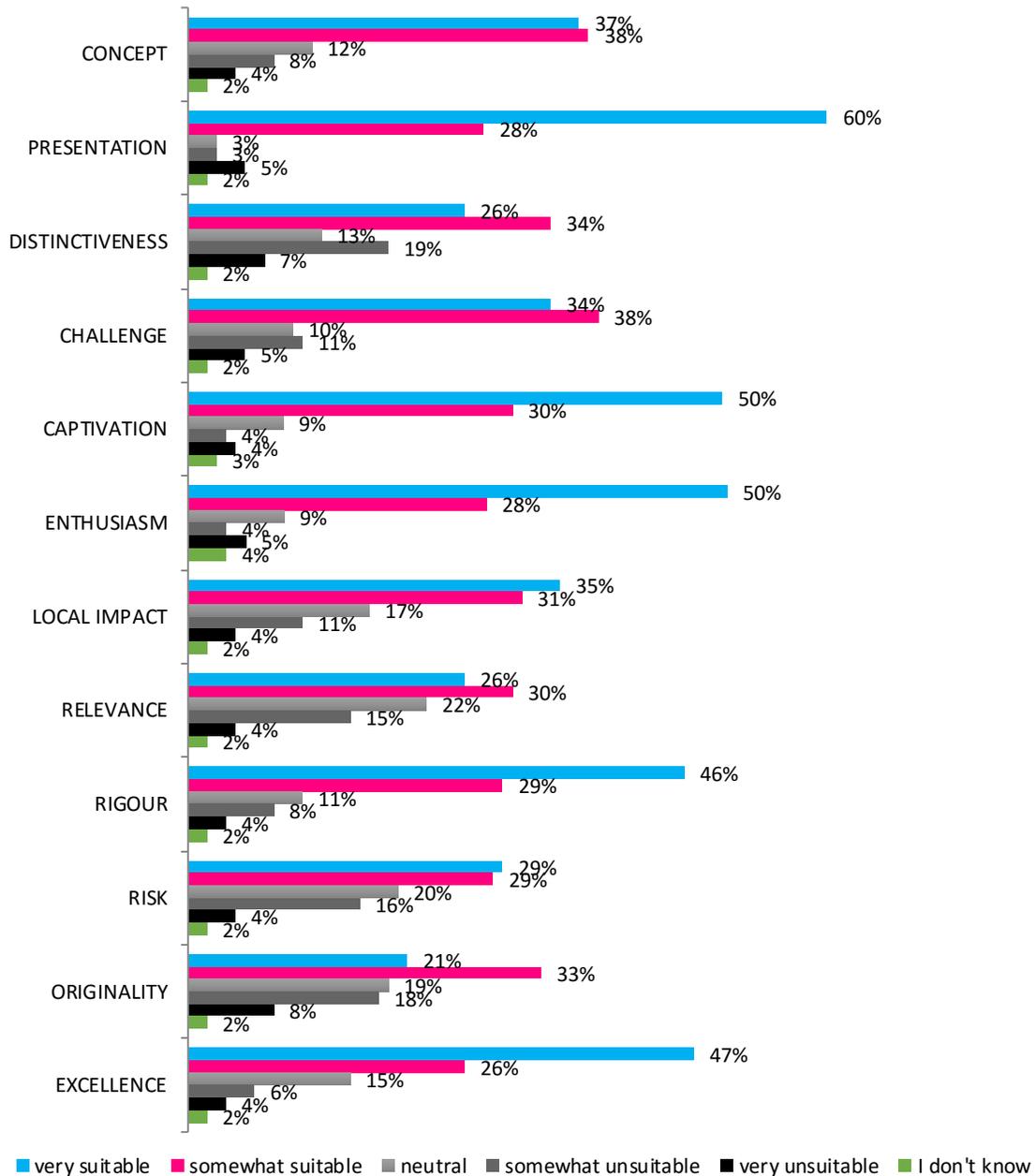
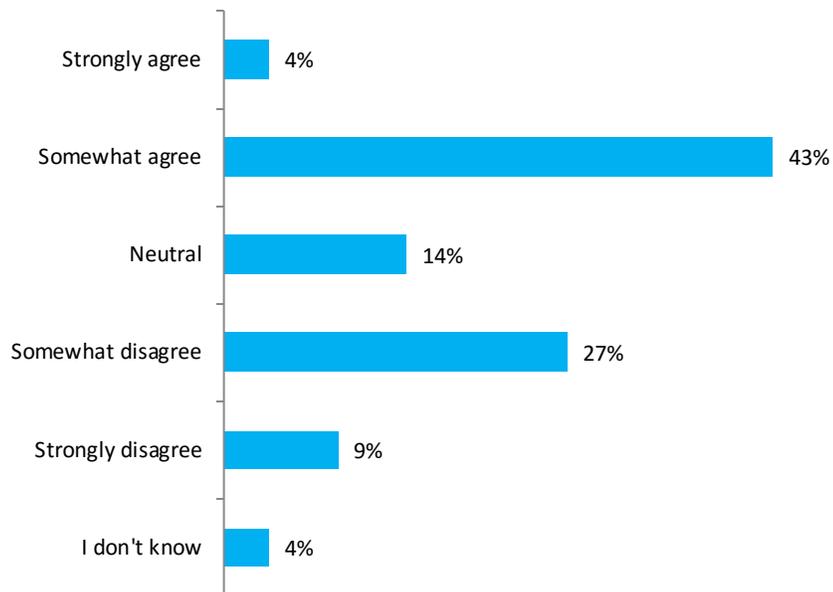


Figure 55: Suitability of individual QM dimensions for assessing quality



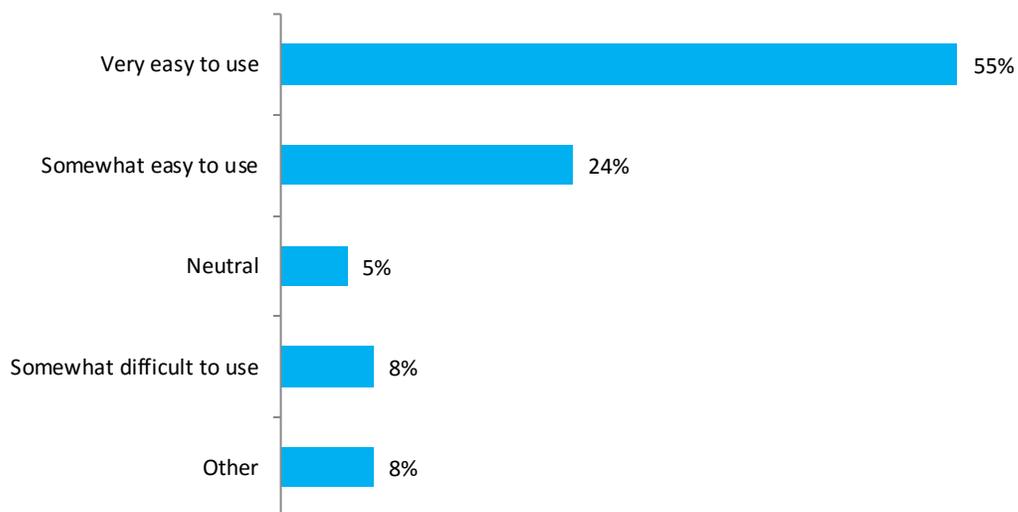
Count = 112, 112, 112, 112, 112, 111, 112, 112, 112, 112, 112, 112

Figure 56: The extent to which peer assessors agreed that QM offers meaningful comparability between different art forms and cultural disciplines



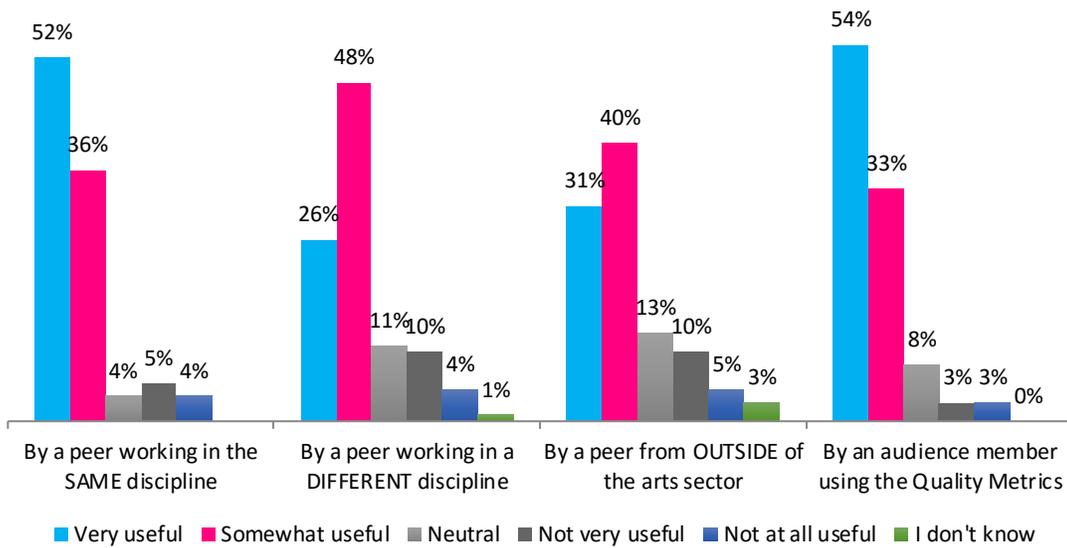
Count = 112

Figure 57: How did peer assessors rate the usability/ease of use of CC platform



Count = 92

Figure 58: Anticipated usefulness of peer and audience assessment to participating organisations



Count = 112, 112, 111, 112

Appendix C: Consultation List

Following is a list of the organisations consulted through the focused interviews, survey of participating organisations, focus groups and email questionnaires. Note: for confidentiality, this list is not exhaustive and does not include individuals consulted, such as peer assessors, in order to ensure anonymity.

Figure 59: List of Organisations Consulted (non-exhaustive)

	Organisation	Region	Discipline
1	Alternative Theatre Company	London	Theatre
2	ARC, Stockton Arts Centre	North East	Combined arts
4	Arnolfini Gallery	South West	Visual arts
5	Arts Council England	National	Other stakeholder
6	Artsadmin	London	Combined arts
7	Audience Agency	National	Other stakeholder
8	Baltic Centre For Contemporary Arts	North East	Visual arts
9	Barbican Centre	London	Music
10	Battersea Arts Centre	London	Theatre
11	Beaford Arts	South West	Combined arts
12	Bedford Creative Arts	East	Combined arts
13	Birmingham Contemporary Music Group	West Midlands	Music
14	Birmingham Museums Trust	West Midlands	Museum
15	Birmingham Royal Ballet	West Midlands	Dance
16	Black Country Living Museum	West Midlands	Museum
17	Blackwell	North West	Visual Arts
18	BookTrust	London	Literature
19	Bournemouth Symphony Orchestra	South West	Music
20	Brighton Dome and Festival	South East	Combined arts
21	Bristol Music Trust	South West	Music
22	Bristol Old Vic And Theatre Royal Trust	South West	Theatre
23	Bristol's Museums, Galleries & Archives (BMGA)	South West	Museum
69	Cambridge Junction	East	Combined arts
24	Candoco Dance Company	London	Dance
25	Carousel	South East	Combined arts
26	Cheshire Dance	North West	Dance
27	Chinese Arts Centre	North West	Visual arts

28	Clod Ensemble	London	Theatre
29	Colchester Mercury Theatre	East	Theatre
30	Collective Encounters	North West	Theatre
31	Coney Ltd.	London	Theatre
32	Contact Theatre	North West	Theatre
33	Corn Exchange (Newbury) Trust	South East	Combined arts
34	Counting What Counts	South East	Other stakeholder
35	Creative Arts East	East	Combined arts
36	Crying Out Loud	London	Combined arts
37	DaDaFest	North West	Combined arts
38	Dance Umbrella Ltd	London	Dance
39	DanceEast	East	Dance
40	Deda	East Midlands	Dance
41	Department of Culture and the Arts, Western Australia	Australia	Other stakeholder
43	Derby Museums	East Midlands	Museum
44	Devon Guild Of Craftsmen	South West	Visual arts
45	Durham County Council	North East	Combined arts
46	Eastern Angles Theatre Company	East	Theatre
47	Eden Arts	North West	Combined arts
48	Emergency Exit Arts	London	Theatre
49	English National Ballet	London	Dance
50	English Touring Theatre	London	Theatre
51	FACT (Foundation For Art & Creative Technology)	North West	Visual arts
52	Fevered Sleep	London	Theatre
53	Free Word	London	Literature
55	Fuel Theatre	London	Theatre
56	Future Everything	North West	Combined arts
57	Gem Arts	North East	Combined arts
58	Gulbenkian Theatre	South East	Combined arts
59	Half Moon Young People's Theatre	London	Theatre
60	Helix Arts	North East	Combined arts
61	Home	North West	Combined arts
62	Ikon Gallery	West Midlands	Visual Arts

63	Improbable	London	Theatre
65	Ironbridge Gorge Museum Trust	West Midlands	Museum
66	J-Night	Yorkshire	Music
67	Jazz North	Yorkshire	Music
68	Junction	Yorkshire	Combined arts
70	Lakeland Arts	North West	Visual arts
71	Lawrence Batley Theatre	Yorkshire	Theatre
72	Leeds Museums and Galleries	Yorkshire	Museum
73	London International Festival of Theatre	London	Theatre
74	Ludus Dance	North West	Dance
75	Mahogany Opera Group	East	Music
76	Manchester City Galleries	North West	Museum
77	Merseyside Dance Initiative	North West	Dance
78	Mind The Gap	Yorkshire	Combined arts
79	More Music	North West	Music
80	Museum of London	London	Museum
81	Music In The Round	Yorkshire	Music
82	New Vic Theatre	West Midlands	Theatre
83	New Wolsey Theatre	East	Theatre
84	Norfolk & Norwich Festival	East	Combined arts
85	North Music Trust	North East	Music
86	Nottingham Playhouse Trust	East Midlands	Theatre
87	Oldham Coliseum Theatre	North West	Theatre
88	Oxford Playhouse	South East	Theatre
89	Peckham Space	London	Visual arts
90	Pentabus Arts	West Midlands	Theatre
91	Pilot Theatre Company	Yorkshire	Theatre
92	Pioneer Theatres	London	Theatre
93	Prism Arts	North West	Combined arts
94	Purple Seven	London	Other Stakeholder
95	Royal Exchange Theatre	North West	Theatre
96	Royal Liverpool Philharmonic Society	North West	Music
97	Royal Shakespeare Company	West Midlands	Theatre

98	Rural Arts North Yorkshire	Yorkshire	Combined arts
99	Serpentine Gallery	London	Visual arts
100	Seven Stories	North East	Literature
101	Siobhan Davies Dance Company	London	Dance
102	Sound and Music	London	Music
103	South East Dance	South East	Dance
104	Southbank Centre	London	Combined arts
105	Spike Island Art Space	South West	Visual arts
106	Spot On Lancashire (Lancashire Rural Touring Network)	North West	Combined Arts
107	Spread the Word	London	Literature
108	Stoke-On-Trent and North Staffordshire Theatre Trust	West Midlands	Theatre
109	Streetwise Opera	London	Music
110	Tangle	South West	Theatre
111	The Albany	London	Combined arts
112	The Anvil	South West	Music
113	The Bike Shed Theatre	South West	Theatre
114	The Brewery, Kendal	North West	Combined arts
115	The Met (Bury Metropolitan Arts Association)	North West	Combined arts
116	The Roundhouse	London	Combined arts
117	The Wordsworth Trust	North West	Literature
118	Tobacco Factory Arts Trust	South West	Theatre
119	Tyne & Wear Archives & Museums	North East	Museum
120	University of Cambridge Museums	East	Museum
121	Watermans (Hounslow Arts Trust)	London	Combined arts
122	Whitworth Art Gallery	North West	Museum
123	York Museums Trust	Yorkshire	Museum

Appendix D: Bibliography

£300k grant sidesteps ACE procurement rules, Liz Hill, Arts Professional (20 November 2015) accessed at <http://www.artsprofessional.co.uk/news/exclusive-ps300k-grant-sidesteps-ace-procurement-rules>

Audience Finder, The Audience Agency, <http://www.audiencefinder.org/>

Culture Counts Results: Henry V, Queen Anne and Wendy & Peter Pan Results Methodological Observations, Loftus, B., Royal Shakespeare Company (2016)

Culture Counts, Counting What Counts, <https://culturecounts.cc/uk/>

Culture Metrics blog, Florack, F., Gilmore, A. & Arvanitis, K. three-way collaboration between a research partner (ICP), a technology partner (Pracsys and Intelligence Agency) and a consortium of cultural organisations including: Cornerhouse and Home (Lead Partner), The Royal Exchange Theatre, Contact Theatre, The Manchester Museum, Manchester Art Gallery, Halle Concerts Society and the Imperial War Museum North, accessed at <http://www.culturemetricsresearch.com/>

Developing participatory metrics, Arts Council England (October 2015)

Digital R&D Fund for the Arts, Arts Council England/Arts and Humanities Research Council/NESTA

Glasgow 2014: Evaluating the Quality of Artist, Peer and Audience Experience, Bunting, C., and Knell, J. (March 2015)

HOME: Quality Metrics – Research & Development Report, Knell, J et al, Digital R&D Fund for the Arts, Arts Council England/Arts and Humanities Research Council/NESTA (November 2015)

Manchester Metrics Pilot: Final Report of Stage One, Knell, J. (May 2014)

Manchester Quality Metrics Pilot 2013: Evaluation Session Report, Soria-Donlan, I., and Gbeleyi, S. (December 2013)

Measuring Quality in the Cultural Sector: The Manchester Metrics Pilot: Findings and Lessons Learned, Bunting, C., and Knell, J. (May 2014)

New study paints picture of arts engagement, Frances Richens, Arts Professional (29 April 2016) accessed at <http://www.artsprofessional.co.uk/news/new-study-paints-picture-arts-engagement>

Public Value Measurement Framework: Issues Paper 2 – Benchmarking, Pracsys, Government of Western Australia Department of Culture and the Arts (August 2011)

Public Value Measurement Framework: Measuring the Quality of the Arts, Government of Western Australia Department of Culture and the Arts (August 2014)

Public Value Measurement Framework: Valuing and Investing in the Arts – Toward a New Approach, Chappell, M. and Knell, J., Government of Western Australia Department of Culture and the Arts (August 2012)



Show Stats, The Audience Agency, <http://www.showstats.org/>

Understanding the Value and Impacts of Cultural Experiences: A Literature Review, Carnwath, J., and Brown, S. (2014)

Using Quality Principles in work for, by and with Children and Young People: Results of a pilot study, National Foundation for Educational Research (NFER) and Shared Intelligence (May 2015)

Appendix E: List of Figures

Figure 1: Respondent organisations' primary discipline	9
Figure 2: Number of events assessed by respondent organisations for national test phase	9
Figure 3: Type of peer assessor	11
Figure 4: Primary cultural discipline or area of expertise of peer assessors	11
Figure 5: Value of QM concept to respondent organisations	12
Figure 6: What respondent organisation thought of particular elements of the QM framework	13
Figure 7: Could respondent organisations anticipate their organisations using the QM dimensions in future	14
Figure 8: Suitability of individual QM dimensions for measuring quality	15
Figure 9: Suitability of QM dimensions for assessing the event attended as judged by type of peer (of the same or different discipline)	16
Figure 10: What respondent organisations thought of particular elements of the QM framework	21
Figure 11: What respondent organisations thought of the triangulation elements	28
Figure 12: Anticipated usefulness of peer and audience assessment to participating organisations	28
Figure 13: Likelihood of respondent organisations using the CC platform in future if cost was not a consideration	34
Figure 14: Helpfulness of each medium of support available for CC platform	38
Figure 15: The administrative burden for coordinating and/or conducting each element of the triangulation	41
Figure 16: Further insight into what respondent organisations thought of the process	42
Figure 17: In what way could the QM framework help the respondent organisations	47
Figure 18: Motivation of respondent organisations to participate in national test phase (frequency)	50
Figure 19: Respondent organisations' primary discipline	64
Figure 20: Respondent organisations' size by number of employees	64
Figure 21: Geographic distribution of respondent organisations	65
Figure 22: Number of events assessed by respondent organisations for the national test phase	65
Figure 23: Level of portfolio investment in respondent organisations per year	66
Figure 24: How many organisation survey respondents were personally also peer assessors	66
Figure 25: Other data collections tools being used by respondent organisations (frequency)	67
Figure 26: Motivation of respondent organisations to participate in national test phase (frequency)	68
Figure 27: Motivation of respondent organisations to participate in national test phase (percentage)	69
Figure 28: Value of QM concept to respondent organisation	70
Figure 29: Value of QM concept to respondent organisations with a diversity focus	70
Figure 30: Value of QM concept to respondent organisations when cross-tabulated with motivation (primary or secondary reason for participating in national test phase) (frequency)	71
Figure 31: Value of QM concept to respondent organisations when cross-tabulated with motivation (primary or secondary reason for participating in national test phase) (percentage)	71
Figure 32: What respondent organisations thought of particular elements of the QM framework	72
Figure 33: In what way could the QM framework help the respondent organisations	72
Figure 34: What respondent organisations thought of the triangulation elements	73

Figure 35: The administrative burden for coordinating and/or conducting each element of the triangulation _____	74
Figure 36: The administrative burden of coordinating and/or conducting each element of the triangulation (Organisations of 50+ permanent staff) _____	75
Figure 37: The administrative burden of coordinating and/or conducting each element of the triangulation (Organisations of 11-50 permanent staff) _____	76
Figure 38: The administrative burden of coordinating and/or conducting each element of the triangulation (Organisations of 1-10 permanent staff) _____	77
Figure 39: Further insight into what respondent organisations thought of the process _____	78
Figure 40: Suitability of individual QM dimensions for measuring quality _____	79
Figure 41: Was help sought from CC for resolving platform problems that organisations had _____	80
Figure 42: Timeliness and quality of the help received in resolving problems with the CC platform* _____	80
Figure 43: Helpfulness of each medium of support available for CC platform _____	81
Figure 44: Likely future intentions of respondent organisations _____	82
Figure 45: Likely future intentions of respondent organisations with a diversity focus _____	82
Figure 46: Respondent organisations appetite for QM under particular future scenarios _____	83
Figure 47: Type of peer assessor _____	84
Figure 48: Primary cultural discipline or area of expertise of peer assessors _____	84
Figure 49: Number of events assessed by all peer assessors _____	85
Figure 50: Location of events assessed by all peer assessors _____	85
Figure 51: Discipline of events assessed by all peer assessors _____	86
Figure 52: Suitability of QM dimensions for assessing the event attended as judged by type of peer (of the same or different discipline) _____	87
Figure 53: Suitability of QM dimensions for assessing the event attended as judged by type of peer (whether AQA or ON peer) _____	87
Figure 54: Suitability of QM dimensions for assessing particular disciplines as judged by all peers (frequency) _____	88
Figure 55: Suitability of individual QM dimensions for assessing quality _____	89
Figure 56: The extent to which peer assessors agreed that QM offers meaningful comparability between different art forms and cultural disciplines _____	90
Figure 57: How did peer assessors rate the usability/ease of use of CC platform _____	90
Figure 58: Anticipated usefulness of peer and audience assessment to participating organisations _____	91
Figure 59: List of Organisations Consulted (non-exhaustive) _____	92



Nordicity

Nordicity is an international arts, cultural and creative industries consultancy specialising in research and evaluation, strategy, policy and economic analysis.

For more information, contact: shignell@nordicity.com

Report by Nordicity. NGL16-09-06

nordicity.com