Spirit of 2012
Data Aggregation Research
November 2023

.fry.

FUNDED BY
SPIRIT OF 2012
INVESTING IN HAPPINESS

Photo by Pietro Jeng on Unsplash

Contents

Background

Spirit of 2012 is the London 2012 Olympic and Paralympic Legacy funder, established with a £47m endowment from the National Lottery Community Fund. They fund projects that help people to be creative, active and connected across the UK in order to improve the wellbeing of individuals and communities. Spirit of 2012 (Spirit) is a spend out trust. Their current strategy sets out an ambition to commit the remaining endowment by the end of 2023 and close in 2026.

As part of committing their remaining endowment, Spirit have recently funded a number of linked research projects aimed at understanding how the culture, sports and events sectors can use, collect, unify and promote data, evaluation and research. This began with the publication of the Spirit of 2012 Inquiry, and has resulted in the commissioning of several interlinked projects that look more specifically into how data and evaluation can be more uniform, better utilised and more influential in these sectors. These projects include 'Connective Tissue' - a consultancy piece aiming to arrive at an ecosystem theory of change for mega events - two scoping exercises, one on the feasibility of a 'UK City of Sport' competition, and one on a 'Data Observatory', which looks to collectivise data, evaluation and research practises from multiple partners and stakeholders for public use. These three projects are being undertaken by multiple partners via separate commissions from Spirit. The 'Data Observatory' scope is being delivered by FRY Creative.

The fourth is this Data Aggregation Research report and the research exercise into aggregating data it documents. Intending to inform and augment the three projects outlined above, it takes three mega events - UNBOXED: Creativity in the UK (UBX), Birmingham 2022 Festival (B22), Coventry City of Culture 2021 (COV) - and works from the ground up to understand if it is possible to centralise, aggregate and then understand the quantitative data from them. As will be discussed, this will involve sourcing data, as the organisations which created these events are now dissolved to varying extents, as is symptomatic of events based projects of this scale.

This research project intends to discover: how to acquire data from these events; how to understand the data based solely upon what is accessed or received; how to aggregate the data so that it can be compared and/or analysed collectively; to anticipate any recommendations for future data collection of this kind, which will make aggregation post-event easier in the future; to suggest opportunities for insight provided by aggregating data in this way.

FRY Creative have been commissioned to undertake the research for this aggregation project. The below outlines: the methodology which was used to undertake this research and test this process; the difficulties encountered whilst accessing, understanding and aggregating the data; and, finally, the opportunities for further analysis and insight encountered in the process. Weaved into the findings are a number of recommendations for future data collection, which would have eased barriers faced when analysing and aggregating post-event.

Methodology

By virtue of being based in the region in which these events took place, FRY had existing and direct access to individuals and organisations who were involved in delivering the listed events. The FRY team - prior to establishing the company - led the evaluation for both UNBOXED and Birmingham 2022 Festival. Evaluators working on these two projects had also been involved in delivering the evaluation for Coventry City of Culture 2021. Subsequently, the methodology which follows, as well as the insights drawn from this process, should be read with this in mind. Via insights presented throughout the report, FRY have endeavoured to be transparent regarding how the above relationships may have affected the ease with which data was accessed or understood.

The project utilises an exploratory methodology as the work is inherently about undertaking a test case for future learning: modelling if/how data can be accessed/understood/aggregated in the context of mega events. Therefore, each of the chapters will discuss the methods we used. In many instances, the methods we used to undertake this work are the insights that respond to research questions set out at the start of the project, attached in appendix one, and referenced below. Chapters, largely drawn from research questions derived by FRY Creative at the inception of this project, and agreed with Spirit of 2012, will be structured as follows:

> **Accessing Data**: how we accessed the data; was the owner clear; can anyone access it; is it open sourced; were data sharing agreements signed; do data sharing agreements limit data use;
>
> **Forms of Data**: how much data was sent; was it raw or in aggregate form; how readable is the data; who/what was the data about; how was it collected;
>
> **Aggregating Data**: how easy is it to merge datasets; what stands in the way; what is the process for aggregating data; is it always possible;
>
> **Further Analysis**: what comparisons do data aggregation facilitate; what do these comparisons tell us; what opportunities for insight are there scaling up this process; how does longitudinal study of data work in this context; is it possible to draw overarching conclusions from aggregation; can collective impact be assessed via this method; what does it mean for place based study.

Upon development of this research we have opted to focus on aggregation of audience data from the three events in question. This is because many of the opportunities/challenges arising from this process, are equally applicable to quantitative data collected via survey with individuals engaging in a project - whether that be as volunteers, participants, staff or otherwise.

Accessing Data

The three projects selected for aggregation were chosen for two principal reasons. The first - and most important - is that there was an element of geographical overlap between them. Birmingham 2022 Festival took place across the West Midlands, UNBOXED delivered one of its major commissions - PoliNations - in Birmingham, and the 2021 City of Culture competition was awarded to Coventry. The second reason - that by proxy of the evaluation work taking place in similar locations, and at similar times, similar groups of people held knowledge about the data. This - as discussed - would aid the processes of aggregation described throughout the report, meaning, perhaps, that more resource could be brought to experiments in aggregation, rather than to navigating barriers to data.

As is symptomatic of events of this scale (and as was the case for UBX, B22 and COV) the delivery vehicles responsible are often temporary. Once the events themselves are delivered, and the evaluation analysis and reporting are underway, the organisations begin the process of dissolution. After the event evaluation is then published, the organisation - as was the case for both UNBOXED and Birmingham 2022 - is fully dissolved. As such, the evaluation reports published, as well as the data collected as part of the evaluation, must be passed on to other bodies to be permanently housed. As the evaluation for Coventry City of Culture was still ongoing at the time of acquisition, data was acquired directly from the Universities contracted by the delivery organisation to evaluate in this instance.[1]

In the context of this aggregation project, the FRY Creative team knew who was now responsible for raw data collected from both UBX and B22. This was because they had been responsible for conducting the handover. For UBX, data was handed to DCMS, as the central body responsible for funding this work. For Birmingham 2022, this data was formally handed over to Culture Central, as they were responsible for delivering some of the legacy outcomes of B22. Data was also sent to the major funders of B22 - fulfilling funding agreements held. As it was known that Cultural Central held the Birmingham 2022 data in this instance, they were contacted for access.

> **Recommendation:** In the future, if the new home of raw data from events is unknown, then the most reliable option - based on the events assessed as part of this project - would be to contact the event funders for access to the data.

Of the three datasets acquired as part of this work, only Cultural Central required a data sharing agreement in order to share the data with FRY Creative. This may have been due to the fact that in the majority of instances the data held was anonymous and therefore did not fall under GDPR restrictions. The only instances where data was not fully anonymised were in open answer questions (which facilitated the opportunity for people to identify themselves), or where full postcode data was still

[1] 'Evaluating Coventry UK City of CUlture 2021' website was launched with all evaluations at the start of November 2023. https://coventry21evaluation.info/

attached to other data points within the dataset - as was the case with B22 data. The data sharing agreement signed for B22 stipulated that data was to be used solely for the purposes of the aggregation project, and for no other purposes. Where data sharing agreements were not required this may have been because FRY Creative were known to the organisations in possession of the data discussed. None of the organisations or the data sharing agreement referred to the consents process that was used at primary data collection.

Forms of Data

Obtaining data often went hand in hand with meeting the people who had collected it. For example, in order for COV to share data, they wanted to understand its use. This was for two purposes: first, to ensure that it was appropriate to share the data and second, to understand the type of data that should be shared. This meeting facilitated a surface level understanding of the data, and crucially the types of data that were to arrive. Meetings were conducted internally at FRY Creative to understand the types of data received from B22 and UBX, as internal members of our team had been responsible for oversight of these projects, and therefore had a strong working knowledge of them.

Figure one sets out the type of data in existence for the three projects. Where labelled 'Y' this data has been shared in raw format with FRY for the purposes of this project. When labelled 'A' this means that we have identified that the data does/did exist, but it has not been shared in raw format with FRY for the purposes of this project. Where data had not been shared, its existence was confirmed by reading published reports and/or methodologies shared with the FRY team. Data was not actively withheld by stakeholders. Rather - as in the case of COV - that the data received via our first request facilitated the analysis necessary for the purposes of this project. As detailed in the methodology section, analysis and aggregation of audience data alone provided sufficient insight into the process by which aggregation can take place. Therefore, no further data was requested from COV.

Data can be shared in multiple formats: in percentage form via reporting, in aggregated form via data tables and/or summaries, or in raw format. Whilst comparison and new insight is possible from data in percentage format or as data tables, aggregation across events is only possible when raw data is shared. In addition, it is only possible to aggregate the same types of data, or data collected in similar ways or via similar methods.

> **Recommendation:** Data should be shared in the rawest possible format - data which has not been cleaned, processed or edited after it has been collected. This will allow all possible aggregation and analysis to take place, whilst also maintaining the integrity of primary data collection. It will also facilitate consistency in the data cleaning process. As will be discussed throughout, despite taking a large amount of time and capacity, then data cleaning should take place centrally to ensure that all data is cleaned/refined in the same way and by the same rules.

One of the first major barriers in aggregating audience data from the three projects was derived from the methods of data collection. The differing methods also hint at one of the more overarching issues faced as part of this process - the intersection between geography and data - which will be discussed later when describing postcode aggregation. Out of the three projects in question, COV intended to have an impact on the smallest geographical location - Coventry - whereas B22 intended to have impact across a region - the West Midlands - and UNBOXED intended to have impact across the four nations which make up the UK. Due to its relative size - COV was therefore able to employ different data collection methods when compared to its geographically more widespread counterparts.

Figure 1. Types of data

(Y = data held, A = data in existence, not held by FRY)

	COV	B2022	UNBOXED
DATA TYPE			
People (un)Involved		Availability	
Population Sample	Y		
Audience	Y	Y	Y
Participant	A	Y	Y
Volunteer	A	Y	
Delivery Staff/Partners	A	Y	Y
Teachers			Y
Student			Y
Project			
Output Monitoring	A	Y	A
Social Media Metrics	A	A	
Media Metrics	A	A	

Based on the data shared for the purposes of this project, audience data collection formed two principal strands as part of COV. The first was similar to that held by UBX and B22 - mixed quantitative and qualitative surveys completed by audience members onsite, during or after their engagement with activity. The second was a general population survey, aimed to be representative of the population of Coventry. In this audience survey, not everyone from the sample engaged with the event, therefore not everyone was an 'audience'. Therefore, at first, we did not aggregate this dataset with that of B22 and UBX, where everyone in the sample engaged as audiences. In short - they were about different stakeholder groups.

As a result, the initial focus of this project's aggregation and analysis has been on the audience datasets collected with people who we know - by proxy of the data collection method - definitely engaged with the work. This is the first of COV's two methods outlined above, as well as B22's and UBX's audience datasets. It is our understanding that these were collected in the following ways:

> **COV:** Surveys distributed digitally and on paper by organisations delivering engagement as part of the event
>
> **UBX:** Surveys delivered by fieldworkers on site at events using representative sampling techniques and surveys distributed digitally by organisations delivering engagement as part of the projects
>
> **B22:** Surveys delivered by fieldworkers on site at events using representative sampling techniques and surveys distributed digitally and on paper by organisations delivering engagement as part of the event

Figure 2. Audience Data

AUDIENCE	COV	B22	UBX
Sample size	9887	5041	5596
Questions (cols)	14	56	48
of which demographic	7	6	6
of which enjoyment	2	2	1
of which sentiment	1	10	22
of which Economic Impact	0	9	1
of which segmentation	1	4	3
of which identifiers	1	6	2

As a major source of audience data for COV was drawn from a general sample of the population, there were a much smaller number of questions in the audience survey delivered at events. Compared to the 56 (B22) and 48 (UBX) questions in its counterparts, the audience survey used for aggregation from COV consisted of 14 questions, as shown in figure two. As will be discussed in the following sections, this somewhat limited the breadth of questions we were able to aggregate across the three projects.

However, in the later stages of this project we were also able to aggregate data from the general population survey conducted by COV. This will be discussed in more depth in the aggregation findings section of the report.

Figure 2 outlines the datasets used for the purposes of this aggreation project. However, other audience data within the respective events is in existence - for example, as highlighted, economic impact analysis as part of UBX. The above represents solely the content of the datasets which were used for aggregation.

In addition to the above, UBX also conducted an economic impact analysis with audiences on site. Whilst this was included in the general audience dataset for B22, UBX data on economic impact was shared in raw form, but in separate excel files to that of the general audience data. This data could also be aggregated with B22 data for comparison. However, due to the way in which it was shared - we are unable to link economic impact data with that of other audience data from UBX. During the evaluation processes for the project, two separate organisations undertook the analysis for the general audience outcomes and the economic impact study. As such the raw data files were split after they had been collected, into two excel files one including audience members' economic impact question responses and one including the rest of audience members question answers. The unique identifiers on each of the raw data files are specific to the data file, and not the dataset as a whole and therefore they cannot be reaggregated with the information currently available. Further work could uncover the original data file with combined datasets.

B22 data was the exception to the rule in terms of how it was shared. All quantitative data collected as part of the whole project was shared as one file. This included the sources for output metrics which were collected from delivery partners on the festival - for example about the total number of participants they engaged - as well as the raw data collected via surveys with volunteers, participants, audiences, staff and freelancers. This enabled the B22 to avoid the issues outlined above. However, it did cause significant problems when it came to manipulating the data - as the file size was so large. The files required significant processing power from computers, and often programmes became slow and/or unresponsive when large amounts of data was being moved. Therefore for the purposes of this project, audience data was extracted and stored in a separate file for analysis.

> **Recommendation:** Where possible, all events should store all data in one master sheet, as a central source of truth. Information can be extracted from it for analysis, but this is the document which should be shared/open sourced due to the possibilities given for aggregation/analysis. This is in line with the prior recommendation regarding sharing data in its rawest possible form.

Aggregating Data

In order to analyse and aggregate data for comparative and longitudinal study, you first need to make the datasets as uniform as possible. The below outlines the process through which we undertook this work, and the resulting barriers and opportunities which came to light as a result.

First, we took an audit of all questions put to audiences in each of the surveys. These questions were then roughly categorised according to the question types outlined in figure two. Due to the fact that the COV had the smallest number of questions, and that we were aiming to aggregate across three events, this was taken as the starting point for looking for continuity across events. Using questions from the COV survey therefore, we searched for equivalents across B22 and UBX. This process led to a focus on 15 questions that we would try and aggregate. Of these 15, there were 8 questions which were present in all datasets, in a form which could be aggregated. This 15 question longlist, and the 8 questions which were successfully aggregated can be found on the 'Audience Question Comparison' page of appendix 2: analysis master. This contains the question wording, as well as the answer options across surveys. A summary is included below as figure three. It splits questions into three areas, each of the question types enabled different exploratory analysis of how aggregation worked: some looked at single answer questions, other continuous scales, other likert scales. The final column of figure three highlights which questions have been successfully aggregated. The results from which can be viewed in appendix 3: Mega Events - Aggregated Data, which is available upon request from Spirit of 2012.

Figure 3. Comparison of Audience Questions

	B2022	UBX	C2021	Merged
DEMOGRAPHICS				
Location	Which of the following best describes where you usually live?	Where do you currently live?	[postcode utilised]	**Y**
Gender	How would you describe your gender?	How would you describe your gender identity?	Please indicate your gender?	**Y**
Ethnicity	Which of the following best describes your ethnic origin or cultural background?	What is your ethnic group?	Please indicate your ethnic origin and cultural background:	**Y**
Nationality	**N/A**	How would you describe your nationality?	**N/A**	**N**
Disability	*Do you identify with any of the following? [if 'Disabled, Neurodiverse or Living with a Long Term Health Condition?]* How would you describe your disability, neurodiversity or long-term health condition?	Do you identify as a D/deaf or disabled person, or have a long-term health condition?	Do you identify as disabled?	**Y**
Sexuality	*Do you identify with any of the following? [if LGBTQIA+]* How would you describe your sexuality?	**N/A**	Sexuality	**N**
Postcode	What is your postcode?	**N/A**	What is your postcode? If you live overseas, what is your country of residence?	**N**
Age	What is your age?	Which of the following groups are you in?	What is your age bracket?	**Y**
Born Overseas	Do you identify with any of the following?	**N/A**	**N/A**	**N**
Mainly Speaking a Language Other than English	Do you identify with any of the following?	**N/A**	**N/A**	**N**

	B2022	UBX	C2021	Merged
VALUE				
Enjoyment	N/A	Please tell us to what extent you experienced these emotions whilst attending [Commission] and/or immediately afterwards? : Enjoyment	I had a good time!	N
Overall Rating	How would you rate your experience overall?	On a scale of 0-10 please rate the [Commission] event/activity you attended?	Please rate the quality of your experience:	Y
General Comments	Do you have any other comments about your experience of the event or Festival 2022?	N/A	Please give us your thoughts about your experience:	N

	B2022	UBX	C2021	Merged
SENTIMENT				
Perceptions of Local Area	Festival 2022 has improved my perception of Birmingham and the West Midlands	Thinking of your experience of [Commission], how has the way you would describe [Commission Location] to someone else changed, if at all?	Did the event have an impact on your perception of Coventry?	Y
Pride in Local Area	It made me feel proud of my local area	How much do you disagree or agree with the following statements in relation to your experience today? : [Commission] made me feel proud to live in [Commission Location]	This experience increases my pride in Coventry as an area.	Y

Highest Common Denominator

As detailed above, as the aim was to aggregate across all three events, we therefore had to begin by using the audience survey with the smallest number of questions. This highlights a central issue at the heart of aggregating datasets. It is an issue which also goes to the heart of why stakeholders undertaking data collection often resist adopting uniform questions, measures and frameworks.

Aggregation of data which has not been designed to be aggregated, or the merging together of questions which are similar but not identical, removes much of the nuances present in that data. Questions would have been developed to provide nuance, resulting in the differences to their counterparts at other events/organisations. Aggregation removes this nuance. In short, in this context, as we wanted to focus on aggregating data across all three projects, we had to use the highest common denominator[2] - the survey with the smallest number of questions of the three.

As we will go on to explore, this principle also applies when attempting to merge the questions/answers/labels themselves. The highest common denominator always dictates how data is merged and analysed. If one question only has two possible answers, but its counterparts have five possible answers each (and we are trying to merge all data into one question) then the five possible answers have to be merged into the two. This is because there is no way of splitting out two answers amongst five more nuanced options. The below outlines this core dilemma at the heart of aggregation, which is displayed through attempts to aggregate various demographic questions.

Aggregating Disability Data

By means of an example, figure four outlines how data on disability was aggregated between surveys. Each event adopted a different approach to asking the question.

In this instance, at B22 it was one option amongst many: to the question ‘Do you identify with any of the following?’ audiences were able to answer ‘Disabled, Neurodiverse or Living with a Long Term Health Condition?’, and give more qualitative information if they wanted to. As not all audiences gave more qualitative information, the information we know for most audiences is if they ticked ‘Disabled, Neurodiverse or Living with a Long Term Health Condition?’ or not. Equally, at COV, audiences could only answer ‘yes’, ‘no’ or ‘Prefer Not To Say’. Therefore in order to aggregate the three surveys you have to lose the nuance established by UBX in allowing people to identify themselves as ‘D/deaf’ as well as ‘a disabled person or person who has a long-term health condition’. The data in aggregate form is forced to become:

[2] Highest Common Denominator in this context is used as a way of describing the method through which you establish commonality between data (common denominator), but with the emphasis on still maximising the use of commonality.

Do you identify as D/deaf, Disabled, Neurodiverse or Living with a Long Term Health Condition?

- Yes
- No
- Prefer not to say

Figure. 4 Disability

Disabled, Neurodiverse or Living with a Long Term Health Condition					
B22		**UBX**		**COV**	
Do you identify with any of the following? [if 'Disabled, Neurodiverse or Living with a Long Term Health Condition?] **How would you describe your disability, neurodiversity or long-term health condition?**		**Do you identify as a D/deaf or disabled person, or have a long-term health condition?**		**Do you identify as disabled?**	
Org. Label	**New Label**	**Org. Label**	**New Label**	**Org. Label**	**New Label**
Yes	Yes	Yes, I am D/deaf	Yes	No	No
No	No	Yes, I am a disabled person or have a long-term health condition	Yes	Yes	Yes
		No	No	Prefer Not To Say	Prefer Not To Say

In the aggregate dataset it is now no longer possible to define between those who are ‘d/Deaf’ or ‘a disabled person or person who has a long-term health condition’ as was possible at the point of collection. The concern is that the more datasets which are added, the more nuance is reduced, as the more uniform answers have to come. However, this problem is core to the concept of aggregation, and not one specific to this context. A method of mitigation is suggested in the below recommendation.

In the above example, the ‘new labels’ which have been applied by aggregating the data, have been derived from the highest common denominator of the options in the three datasets we are working with. However, as established, this can be reductive to the data. A better model would be for a centralised authority to stipulate uniform labels. If a dataset is to be aggregated into a central pool, then it must:

1. Collect data using these labels
2. Collect data in a way which can be aggregated into these labels at a later date.

This will produce a threshold at which data is accepted into the aggregation. If the data doesn't meet the adequate level of granularity to be subsumed by the labels set out by the central authority, then it will not be aggregated. The also works to set a precedence for the level of nuance questions will hold.

> **Recommendation:** A centralised authority should be responsible for establishing uniform questions and uniform answers/labels. This should be open sourced for events to use as a guide. The incentive to the organisation would be that collecting data in this way will facilitate easier and more robust comparison against industry standards. As will be discussed, in order for aggregation to work, those collecting data must be invested in the ability to aggregate post event.

Aggregating Ethnicity Data

In addition, aggregation in and of itself may be problematic. This most prominently comes to light when thinking about demographic data. The below example on ethnicity extrapolates the point. All three projects used a similar level of granularity with ethnicity data, as shown in figure five below.

However, there are underlying questions about whether ethnicity data (or any data pertaining to identity) should be changed from the original way in which people identify. For example, if someone self identifies as 'Pakistani' does this mean they can be aggregated as people who identify as 'Asian or Asian British: Pakistani'. Despite identifying as 'Pakistani' that person might not identify as 'Asian', and in this respect, to what extent are you doing a disservice to asking someone about their identity in the first place, if you then remove nuance from their answer.

A slightly different angle on the same problem comes in reporting on ethnicity data. Despite the above lists being used as options on quantitative surveys by organisations, in reports, in particular in summaries/press releases, people's ethnicities are grouped. For example, as in the case of B22, Asian or Asian British, Black or Black British, Middle Eastern or Arabic, Mixed Heritage, White, Other were aggregated for the final report. We have also presented ethnicity data in the same way for this report - as it enables more readily understandable graphs.. As above, there are arguments that it is inherently reductive to ask individuals (in this instance audiences) to identify themselves in a nuanced way, to then remove that nuance through aggregation after the fact.

The mitigating factor in this instance is that 'groups' of answers for B22 were prefaced with the way in which they would later be aggregated. For example: 'Asian or Asian British: Pakistani' as an answer option. When someone ticks this they are identifying themselves as both 'Pakistani' and as 'Asian or Asian British' - therefore the aggregation of them into a group included in their answer is more permissible.

Finally, as will be discussed in the 'survey transcripts' section below - the list of possible responses from UBX caused issues for aggregation. This list of answer options was derived from the raw data, not from a survey transcript. In the labels derived from the raw data, there was no group preface. This returns us to a similar problem - if someone identifies their ethnicity solely as 'African' - as it appeared to be in the UBX raw data - then they cannot be aggregated with people who Identify as 'Black or Black British: African'.

In the instance of the UBX labels, we had the advantage of close contact with the people who collected this information. In this instance, the answer options present in the raw data did not accurately reflect the questions as they were asked to members of the public. Upon discussion, and upon seeing a survey transcript, the answer option 'African' in the survey was prefaced with 'Black or Black British. In the process of data cleaning, the preface had been removed from the label.

Figure. 5 Ethnicity

B22	UBX	COV
Org. Label	**Org. Label**	**Org. Label**
Asian or Asian British: Chinese	African	Any Other Asian Background
Asian or Asian British: Indian	Arab	Any Other Black Background
Asian or Asian British: Other Asian background	Bangladeshi	Any Other Ethnic Group
Asian or Asian British: Pakistani	Caribbean	Any Other Mixed Background
Black or Black British: African	Chinese	Any Other White Background
Black or Black British: Caribbean	English/Welsh/Scottish/Northern Irish/British	Arab
Black or Black British: Other Black/African/Caribbean background	Gypsy or Irish Traveller	Asian or Asian British
Middle Eastern / Arab background	Indian	Asian or Asian British - Bangladeshi
Mixed: Other mixed/multiple ethnic background	Irish	Asian or Asian British - Chinese
Mixed: White and Asian	Other Asian background	Asian or Asian British - Indian
Mixed: White and Black African	Other Black/African/Caribbean background	Asian or Asian British - Pakistani
Mixed: White and Black Caribbean	Other ethnic group	Black or Black British - African
White: English/Welsh/Scottish/Northern Irish/British	Other mixed/multiple ethnic background	Black or Black British - Caribbean
White: Other White background	Other White background	Mixed or Multiple - White and Asian
Other: All other backgrounds [specify]	Pakistani	Mixed or Multiple - White and Black African
Prefer not to say	Prefer not to say	Mixed or Multiple - White and Black Caribbean
	Roma	Mixed or multiple ethnic group
	White and Asian	Other (please specify)
	White and Black African	Prefer Not To Say
	White and Black Caribbean	White - British
		White - Gypsy or Irish Traveller
		White - Irish

As noted elsewhere in this report, this learning highlights two things: first, that access to people who collected the data is crucial for successful aggregation. Without them, the time taken to discover the above may have taken significantly longer. In order for this communication to occur, the original data owners need to be invested in aggregation being successful. Second, as with the recommendation detailed below in the 'survey transcripts' section, access to a survey transcript along with the raw data would have prevented this problem. It also shows how, by sharing data in the rawest possible form, issues could be prevented.

> **Recommendation:** Access to stakeholders who had an involvement in collecting data is essential for aggregation to take place. These stakeholders must therefore be motivated for aggregation to succeed.

Circling back to grouping answer options into a question with the highest common denominator, grouping ethnicities may also then facilitate loss of meaning. For example, in certain locations there may be particularly strong representation from an ethnicity group which may not be as strongly represented in other geographical areas. Events taking place in a specific location, such as a city, may therefore wish to add these options to their list, to empower these groups. Does aggregating these groups then lose the nuances of engagement for that specific event, in that specific location. For example, a city with a strong Irish population includes 'White: Irish' on their answer options list, but this is merged with 'White: European' or 'White: All other white backgrounds' in aggregation. Has a crucial piece of information been lost about engagement which took place with a specific group of people, in a specific place.

One major advantage and opportunity with aggregating data on ethnicity is around 'Other: All other backgrounds [specify]' answer option. This option, present on both B22 and COV surveys allows people to self identify their ethnicity. By gaining access to multiple records of how people choose to self-identify, across different events, would facilitate insight into the ethnic groups which are more frequently likely to self-identify. If people are frequently forced to write in their own ethnic group, then this could in turn provide the evidence needed to establish this as a new uniform label to be added to a centralised list. In turn, this may help people from this group feel more represented and seen in spaces where they have traditionally been minoritised, whilst also future proofing the processes of data aggregation. Whilst this is an ideal, it is also time consuming due to coding ramifications.

Aggregating Geographical Data

Geography provides an incredibly important opportunity for data aggregation, as will be detailed in the opportunities for further analysis section of this report. Data aggregation on a large scale could facilitate segmentation of engagement via city, region, nation.

As detailed in 'Audience Question - Aggregation Analysis' sheet of appendix 2, aggregation of audiences 'home location'/geography resulted in three possible answers: 'Elsewhere in the UK', 'Commision location' and 'Overseas'. This inherently ties the audience's response to the event that they are experiencing. In addition, if they are not from the place in which the commission is taking place, it becomes impossible to know where in the UK the individual is from.

Both B22 and COV data raw data was shared with postcodes. This means that we are able to identify where the individual's home location is, with a high level of specificity. UBX data was shared without postcodes. As above, without access to the UBX team, and/or the transcripts for the surveys, the raw data alone does not make it clear if postcode data were collected. Upon conversation with the UBX team, postcodes were collected, but not included in the raw data due to GDPR concerns.

Collecting full postcodes from audiences is often a concern on the part of evaluators and event organisers alike. This is because, if sample sets are small, and if postcodes only cover a small number of homes, then stakeholders risk individuals becoming identifiable in the dataset. This issue is somewhat negated when working with sample sizes in the 1000s, but nonetheless hesitancy does prevail.

Postcodes offer a huge number of opportunities for an aggregated dataset. As mentioned above, it means that we can tie an audience's experience to being from a particular city. When we aggregate data from multiple events happening in this city, we have the opportunity to then begin comparing outcomes on a city level, taking into account data from multiple events.

Moreover, postcodes provide us with a number of other segmentation opportunities for data. Full postcode enables us to link an audience survey to:

- The council ward they belong to and therefore their local authority
- The Indices of Multiple Deprivation rating of that location
- The LSOA and MSOA areas relating to that postcode, and any other data relating to this information

In short, access to postcodes that are linked to audience surveys opens the door to a number of other external sources of information, which provide greater segmentation opportunities for the aggregated dataset. In addition, by linking the audience returns to a geographical location, we are able to better

identify their exact relationship to the event they are engaging with, as well as gain an overarching view of engagement in that area.

> **Recommendation:** When data is collected it is always linked to postcode information, and where this is not given by the participant an alternative indication of where they live, for example a town/city/region, is provided.

Aggregating Age Data

Aggregation of age data was problematic due to the way in which answer options were available to audiences. B22 collected continuous data, asking audiences for their exact age. This was in contrast to UBX which asked audiences where they fit within 10 age brackets, and COV, which asked people to identify within 5. This becomes a problem when the brackets overlap with each other. This is shown in figure six below.

As COV had the lowest number of answer options, aggregation began from this point. However, because UBX had grouped ages as 35 to 44 / 45 to 54 / 55 to 64 we faced an issue when relabelling its data according to COV's labels. The UBX grouping 45 to 54 straddled two groups from the COV dataset: 35 to 49 and 50 to 64. As such there was no way of identifying which of the UBX audience surveys should be recategorised into 35 to 49 and which into 50 to 64. Therefore in order to complete aggregation of age, the only option we had was to split all audience data returns within the 45 to 54 category in half, and then split them equally, but arbitrarily, amongst 35 to 49 and 50 to 64. This was ultimately an unsuccessful attempt at aggregation as the 45 - 54 split between groups occurred randomly. The purpose of pursuing aggregation for this demographic question was to test the limits of what is possible and to display the importance of alignment at primary data collection point. As such all findings presented regarding age should be read with this in mind.

Evidently, the way to solve this problem would be to collect continuous data on age. However, when data is collected in a fieldwork setting, with interviewers conducting interviews, there is much feedback that it is too intrusive for some audiences to disclose their exact age when completing surveys. In spite of this, for aggregation around age data to work successfully there are only two options: all parties use the same age brackets, or all parties collect continuous data. To avoid the intrusive nature of demographic data questions in fieldwork contexts, many organisations ask the respondent to complete this part of the survey independently.

> **Recommendation:** Age data is always collective in continuous format, with the option that people don't have to answer the question if they don't want to. Demographic data is collected independently in a fieldwork setting.

Figure 6. Age

Age					
B22		**UBX**		**COV**	
Org. Label	**New Label**	**Org. Label**	**New Label**	**Org. Label**	**New Label**
CONT.	0 - 19	Under 18	0 - 19	0 - 19	0 - 19
	20 - 34	18 to 24	20 - 34	20 - 34	20 - 34
	35 - 49	25 to 34	20 - 34	35 - 49	35 - 49
	50 - 64	35 to 44	35 - 49	50 - 64	50 - 64
	65 +	45 to 54 (50%)	35 - 49	65 +	65 +
	Prefer Not To Say	45 to 54 (50%)	50 - 64	Prefer Not To Say	Prefer Not To Say
		55 to 64	50 - 64		
		65 to 74	65 +		
		75 to 84	65 +		
		85 or older	65 +		
		Prefer not to say	Prefer Not To Say		

Survey Transcripts

When data was shared with FRY, it was provided to us in raw form. None of the data provided by any of the events came with the template surveys that were used to collect the information. This resulted in a difficulty understanding how questions were asked. For example, could a participant only pick one option or multiple, was there survey logic in operation, were all questions optional. It also meant a lack of certainty in developing an exhaustive answer options list. With the information available, the only way to gather a list of the different ways people were able to answer, was to run a =UNIQUE function on the raw data column, to draw out all unique returns for that question. In short, this meant that if no one in the dataset selected a particular answer, then we have no way of knowing it was an option in the first place. This would be solved via access to survey transcripts.

> **Recommendation:** When raw quantitative data is shared, do so with a copy of the survey, as it was provided to the respondent who completed it. This recommendation would also provide clarity in the consents used whilst gathering information.

Selecting the Sample - Weighting and Collection Method

We also faced a number of barriers in using the samples we were provided - this provoked questions about what should be included and excluded from analysis.

Primarily, raw data for UBX came with an indication of how the response ought to be weighted within the overall sample. This in short means a straight analysis of the raw data as it was shared would result in differing percentages to those presented in final reports - this is because analysis would take place without taking into account sample weighting. However, the percentages produced for reporting could be replicated through analysis of the raw data, by taking into account the sample weighting linked to each of the audience survey returns. When looking to aggregate this data with other datasets from different events, should we take into consideration weighting - allowing more precedent to certain returns in the aggregated dataset?

We landed that in aggregation it would be better to keep the sample unweighted. This was for two reasons. First, that the other datasets we had been using did not come with suggested weightings. Second, once the data has been aggregated in fully raw form, it is then possible to apply new weightings to all data, post-aggregation.

In a couple of instances, datasets provided had not been adequately cleaned. This came to light during the process of aggregating age brackets. In the B22 data, around 100 returns were effectively under age 16. This was in spite of the fact that the methodology for the evaluation suggested that no data would be collected with under 14s. Some of the audience data returns present in the final dataset listed their age as '1'. This was clearly a typo, however it begs the question as to whether this data ought to be removed, or whether it should be recategorised as 'undisclosed' or 'Prefer not to say.' It therefore becomes difficult to know where to draw the line between presuming the answers under 16 were typos as opposed to presuming that it was a mistake in the fieldwork sampling, and someone was sampled who was indeed under the age of 14.

This line of questioning points to the level of decision making involved in the cleaning process, if this cleaning occurs prior to data being aggregated, then this will skew the aggregated dataset. For example, one organisation may have opted to delete these audience returns prior to aggregation, one organisation may have opted to change all responses identifying as under 14 to 'undisclosed' age, some may have left them as they were when the survey was returned. If all data is provided to the aggregator in completely raw format, then these rules can be established centrally, then applied, to ensure that all data is playing an equal role in the aggregated set.

In addition, when provided with raw data, the analysis conducted from the raw data may produce different % outputs to that listed in the evaluation reports from the events and/or organisers themselves. This is because the decision making process above has altered the sample used to derive final results. This may result in disputed outcomes/impacts. This is further compounded by the fact that methodology statements often do not provide granular detail as to how audience samples were cleaned.

The most extreme version of the what to include/what not to include dilemma is shown in analysis of B22 data. Much of the data presented in the B22 overall evaluation report opts to solely use the fieldwork sample of audience data. This amounts to around half of the audience survey returns present in the whole dataset. The other half of the returns were collected either digitally or via the delivery partners producing work conducting data collection outside of their external fieldworker allocations. In contrast to the B22 final report, and to test the impact of doing so, we have used the full sample of audience surveys provided. Equally, if we were to only have knowledge of the raw data, and no external knowledge of the reporting, then this would be the natural option to take.

By means of example, in processing the data for this project, we conducted an analysis of B22 audience data. Upon analysing this raw data for the purposes of this project, we found that 5% of B22 audiences identify as 'neurodiverse, disabled or living with a long term health condition'. In B22's final report this figure is listed as 7%. This is because the final report published by B22 had utilised only the fieldworker-collected audience surveys as part of its final sample, whereas for the aggregated analysis we had utilised all audience surveys - including those collected by projects, and those collected online. The discrepancy in % of people identifying as 'neurodiverse, disabled or living with a long term health condition' is symptomatic of how % will differ slightly across the board, due to different returns being used/removed from analysis. In this particular context, the difference may well be because people who identify in this way were less likely to opt in to fill out audience surveys.

The ability to label all survey returns according to whether they were delivered by fieldworks or whether they were self-completes would also be a valuable asset to data aggregation going forward. It would enable us to understand how different research approaches bias audiences from certain backgrounds. It would provide strong, robust and representative samples that would give evidence and leverage that more funding should be available for events based organisations procuring independent, third party fieldworkers to use systematic sampling approaches.

> **Recommendation:** All raw data provided should be coded according to how it was collected: fieldworkers, vs self-complete in person, vs self-complete online.

Aggregating Scales

Another minor challenge faced when aggregating scale based questions was how to merge agreement questions with different scales. This principally comes to light when discussing the % of audiences that 'agreed' with a statement. Where the options were an agreement scale, this is made simpler, but in the case of B22, this was a scale of 100 points between 0 and 1. In this instance we have used scores of above 0.62 to be defined as 'agreement', as is detailed in figure 7.

Figure 7. Scales vs Agreement

B22		**UBX**		**COV**	
Org. Label	**New Label**	**Org. Label**	**New Label**	**Org. Label**	**New Label**
0.81 - 1	Strongly Agree	Strongly agree	Strongly Agree	Strongly Agree	Strongly Agree
0.62 - 0.8	Agree	Agree	Agree	Agree	Agree
0.51 - 0.62	Not Sure	Neither agree nor disagree	Not Sure	Not Sure	Not Sure
0.21 - 0.5	Disagree	Disagree	Disagree	Disagree	Disagree
0 - 0.2	Strongly Disagree	Strongly disagree	Strongly Disagree	Strongly Disagree	Strongly Disagree
		N/A or Don't Know	Not Sure		

Aggregation Findings & Opportunities for Further Study

The final section of this report looks at the insights generated as part of aggregating the data, using them as a springboard to understand the opportunities which data aggregation can provide for insight in general. It highlights how, via data aggregation, we can make better use of quantitative data to understand the cultural sector across the UK, as well as to articulate the narrative of the impact that events such as those discussed can have. While often complex and problematic, the opportunities for aggregation, and the methods of making aggregation easier as identified above, would provide significant opportunity for deepening understanding of: place based investment in events; longitudinal impact of cultural and sporting programmes on audiences; understanding and engaging those who are most marginalised in society; fostering better data-driven decision making for funding. This list is not exhaustive.

As part of this project five datasets were successfully aggregated into the Mega Events Dataset as attached in appendix 3. These were the three audience surveys delivered with people who engaged with the work at B22, UBX and COV, as discussed in the opening chapters of this report. In the latter stages of aggregation we also incorporated data from waves two and three of the population survey delivered for COV, which enabled analysis of two additional sentiment outcomes: wellbeing and community connection. A summary of the data points aggregated across datasets can be found below in figure 8.

Figure 8. Aggregated Datasets by Data points

Project	B22	UBX	COV	COV	COV
Type	Aud	Aud	Aud	Pop2	Pop3[3]
Unique Identifier	Y	Y	Y	Y	Y
Gender	Y	Y	Y	Y	Y
Location	Y	Y	Y	Y	Y
Age	Y	Y	Y	Y	Y
Disability	Y	Y	Y	Y	Y
Ethnicity	Y	Y	Y	Y	Y
Overall Rating	Y	Y	Y		
Perceptions of Local Area	Y	Y	Y	Y	Y
Pride in Local Area	Y	Y	Y	Y	Y
Wellbeing & Happiness	Y	Y		Y	Y
Community Connection	Y	Y		Y	Y

The below figures 9 - 13 plot out each of the demographic questions aggregated above - highlighted in light orange in figure 8. They assess the extent to which certain demographics were more or less likely to agree with the sentiments we were able to aggregate -

[3] Pop2 was collected in Summer of 2022, and Pop3 in December 2022

presented in figure 8 in darker orange.

In the figures 9 - 13 below, the X-axis at 0 represents the average answer given across the population. Using figure 9 as an example, this means that Asian or Asian British Audiences were 2.8% more likely to agree that the event they experienced had a positive effect on their wellbeing, whereas audiences who identified as 'All other backgrounds' or who adopted to self-identify, were 2.8% less likely to agree with the same statement. A breakdown of the sample sizes used for each of these figures can be found in the 'Analysis' tab of 'Mega Events - Audience Data' attached as appendix 3.

Below we have briefly outlined some of the primary findings from this aggregation, but it is recommended that further analysis take place to clarify findings. For example, it would be useful to explore weighting the dataset so that events hold equal precedence, in order to check if results maintain. Again, these findings are brief and non-exhaustive, as the main purpose of this project was to cover the potential opportunities for analysis and insight generated by aggregation.

Aggregation Findings

One of the most pervasive findings across these figures is the extent to which people who opted to not disclose demographic information were much less likely to agree with sentiment based questions. This was near universal across all sentiment questions and demographics, with the exception of wellbeing in figure 13. Audiences who opted to not disclose their demographic information (who picked 'Prefer not to say' in response to demographic questions) were much less likely to feel:

- an increased sense of local pride;
- an improved perception of the local area;
- an increased sense of wellbeing;
- an increase sense of community connectedness.

Further exploration in this area could uncover if it was the same audiences across demographic questions who continually opted not to disclose identity based information or if audience members opted not to disclose one of their demographic identities. Moreover, further research into the reasons why people opt not to disclose this information would be needed in order to understand how to combat the disproportionate effect seen on this/these audience group/s.

As shown in figure 9, which plots aggregated ethnicity data against the sentiment questions, across the board those from more traditionally minoritised ethnic backgrounds, were more likely to report positive experiences of all sentiment questions, when compared to audiences from white backgrounds. The greatest extent to which opinion varied based on ethnic identity was on 'perceptions'. Discrepancy between answers from different ethnic identities was smallest on sentiment questions around overall levels of enjoyment.

Figure 10 plots gender data against sentiment. The main takeaway here was that women were more likely to agree with all sentiments than men. Sentiments varied for people who identified as non-binary or who self-described - they were generally more likely to identify a greater sense of community connection and wellbeing, and less likely to identify an increase to local pride. Wellbeing saw the biggest discrepancy when looking at how gender identity affected results. Men were 2.6% less likely to agree that their wellbeing had been positively affected, women were 1.9% more likely to agree, and non-binary people 8.5% more likely to agree that their wellbeing had been positively affected.

Figure 11 displays age. The general trend looking across all sentiment questions is that 20 - 49 year olds were more likely to report a positive experience (across sentiments) of events when compared to those aged 50+. Those aged 19 and under were less uniform across dimensions: they were 8.3% more likely to agree that their perceptions had been positively affected, but 4.9% less likely to identify that their wellbeing had improved as a result. As discussed, findings regarding age should at best be taken as speculative and at worst dismissed completely as aggregation of UBX data was ultimately randomly distributed for one age bracket.

Due to restrictions on the data points held about geographical location in the datasets - the main geographical comparison that was possible was on those who were from the location that the event took place in compared to those who were not from that location. Perhaps surprisingly, those who were from ‘elsewhere in the UK’ attending events, were more likely to agree that events had improved their wellbeing, and slightly more likely to say that it improved their wellbeing. Perhaps unsurprisingly, those from ‘elsewhere in the UK’ were 4.3% more likely to say it improved their perceptions of a place, and 14.5% less likely to say that it increased their sense of local pride.

Finally, figure 13 sets out the difference in results between those who identified that they were d/Deaf, neurodiverse, disabled, or living with a long term health condition, and those who did not. Those who identified as the former, were less likely to feel an increased sense of local pride, less likely to have increased perceptions, and less likely to feel more community connection. However, they were 3.3% more likely to identify that the events had a positive impact on their happiness/wellbeing.

Figure 9. Aggregated Ethnicity Data

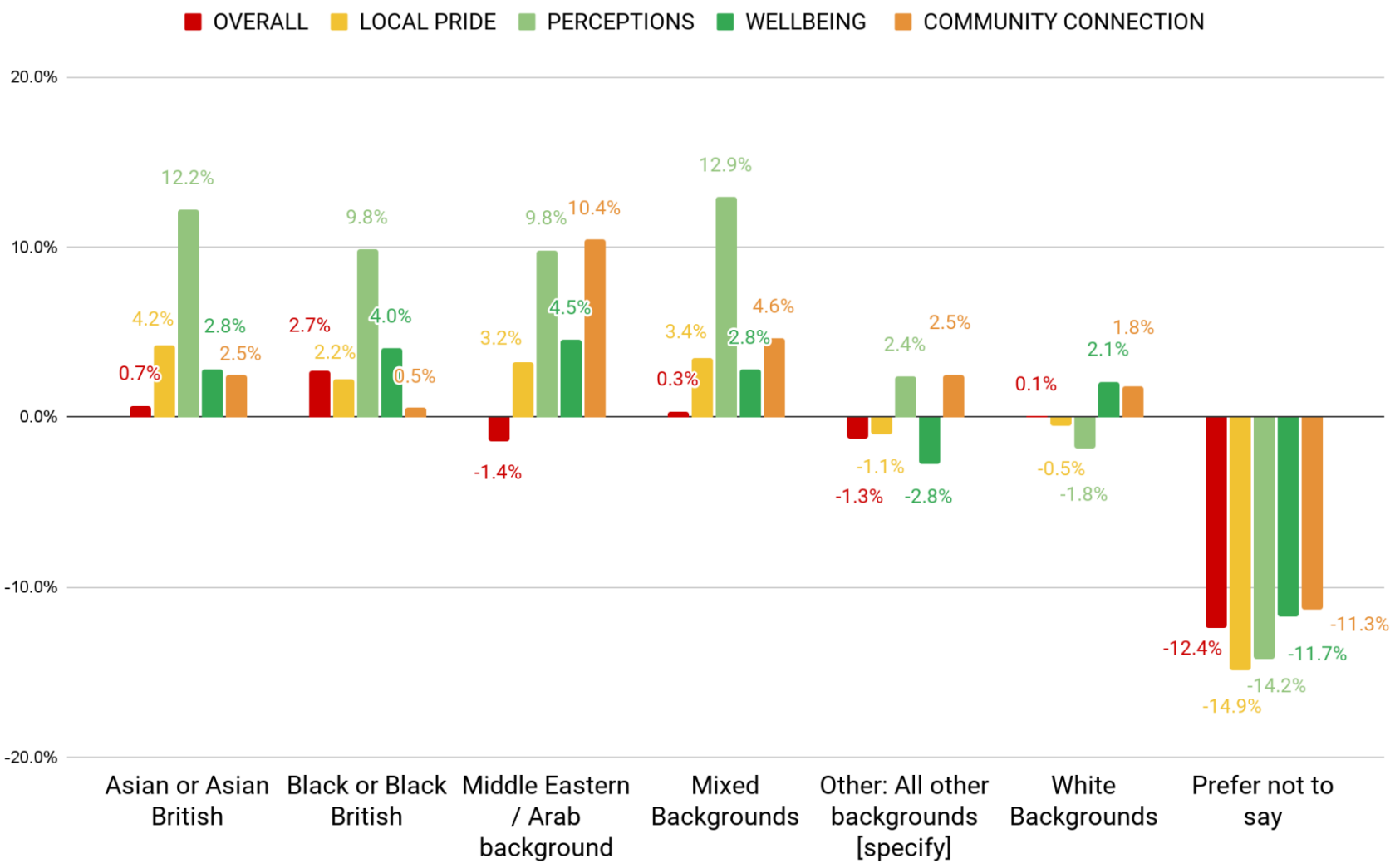


Figure 10. Aggregated Gender Data

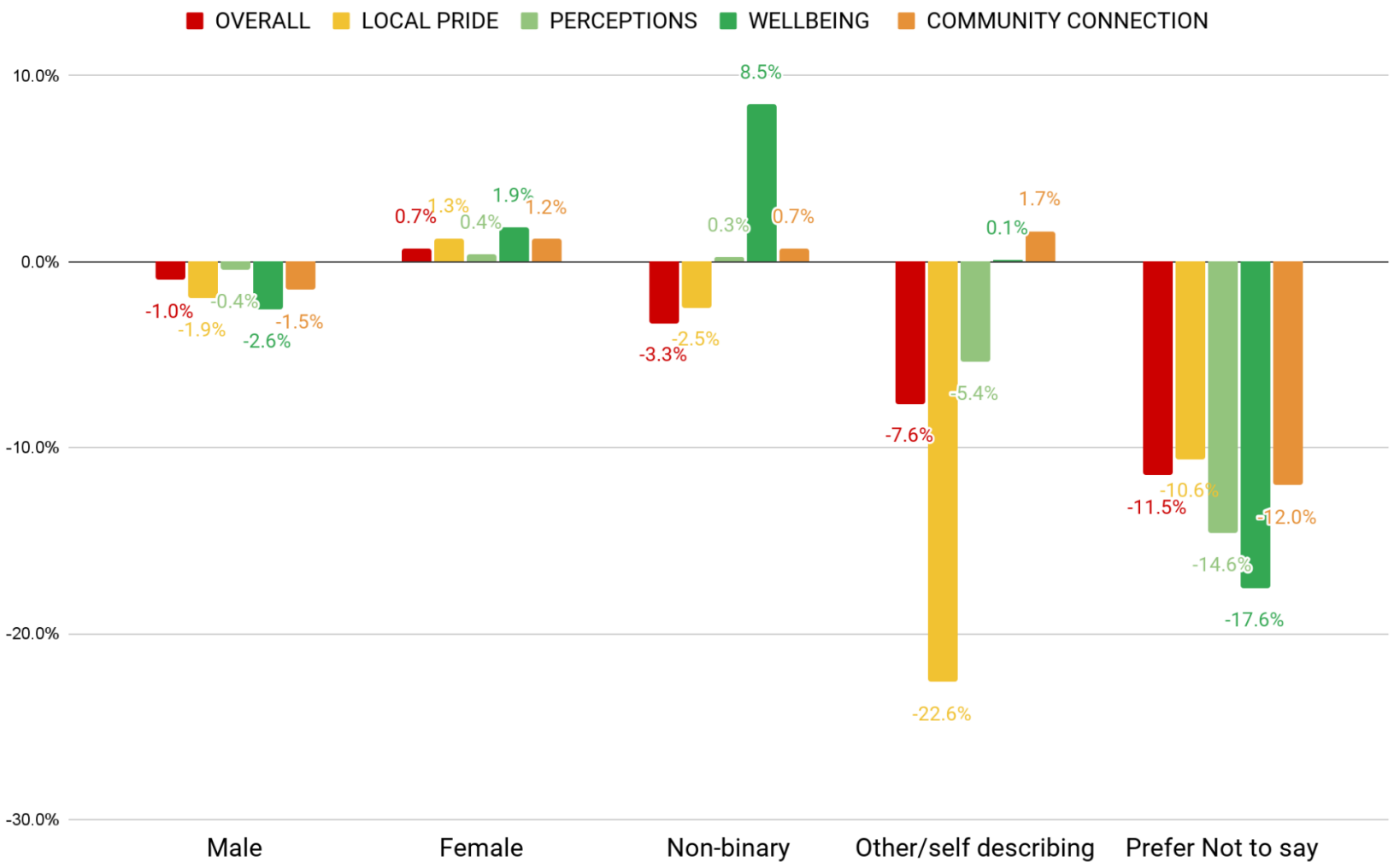


Figure 11. Aggregated Age Data

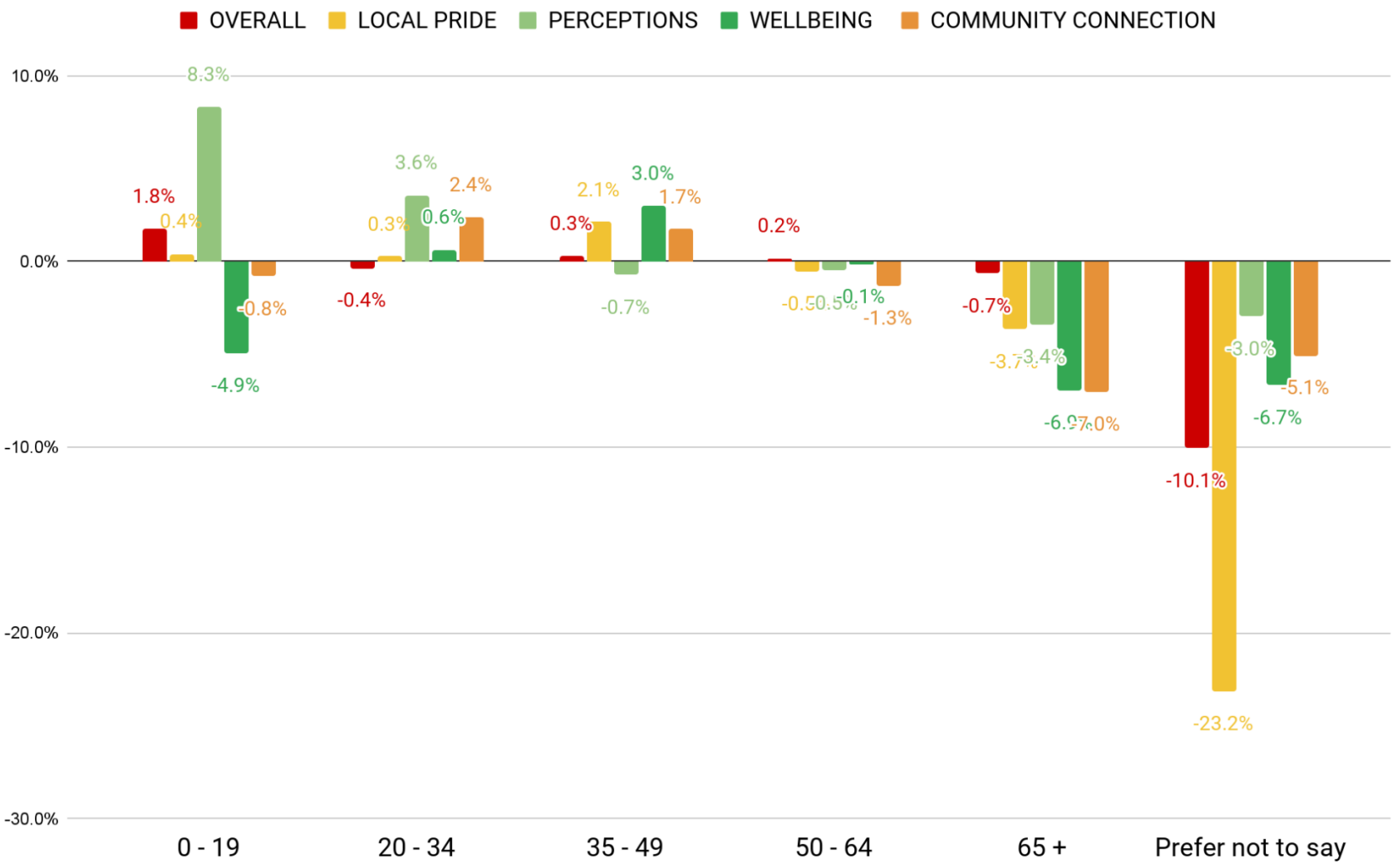


Figure 12. Aggregated Location Data

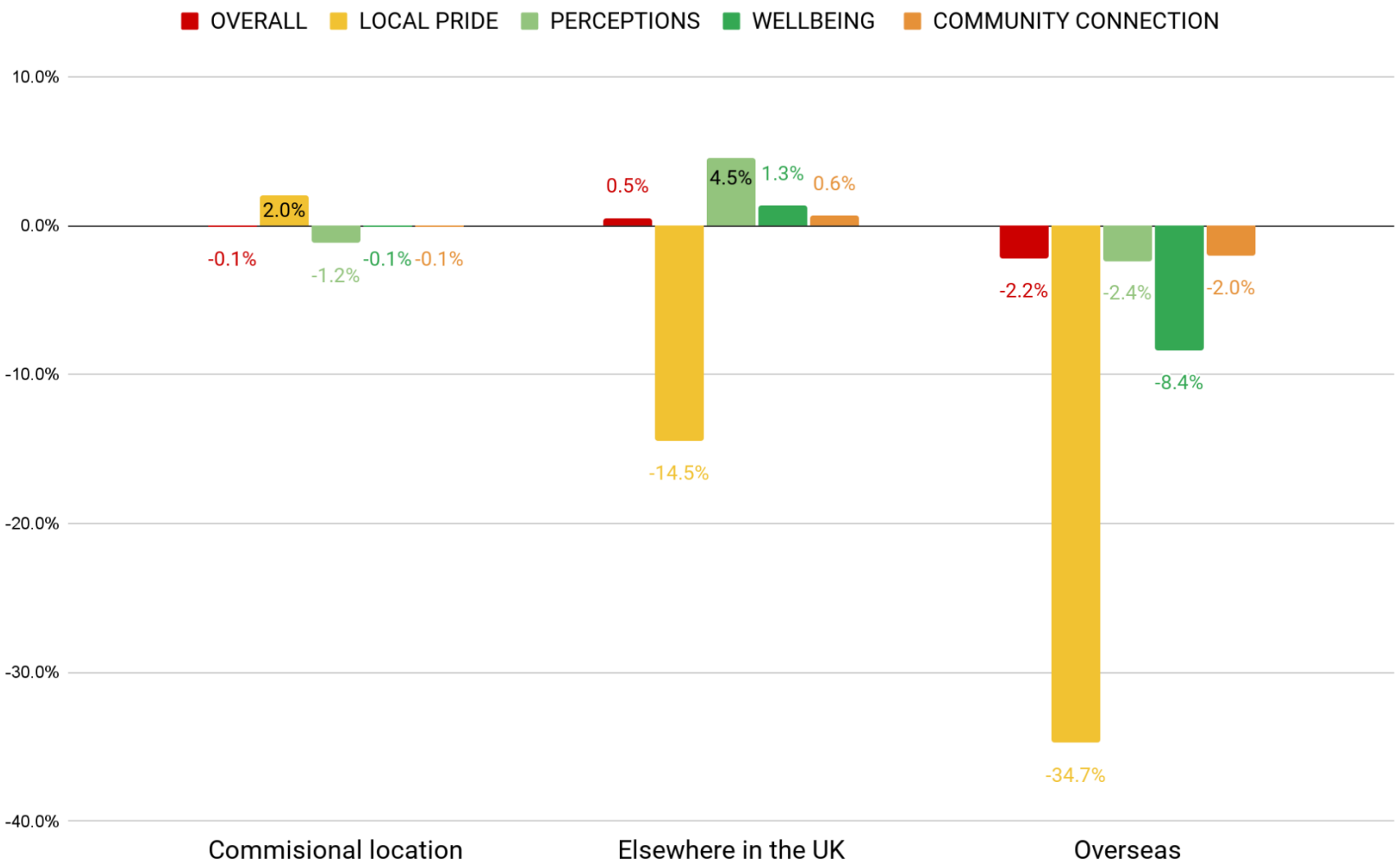
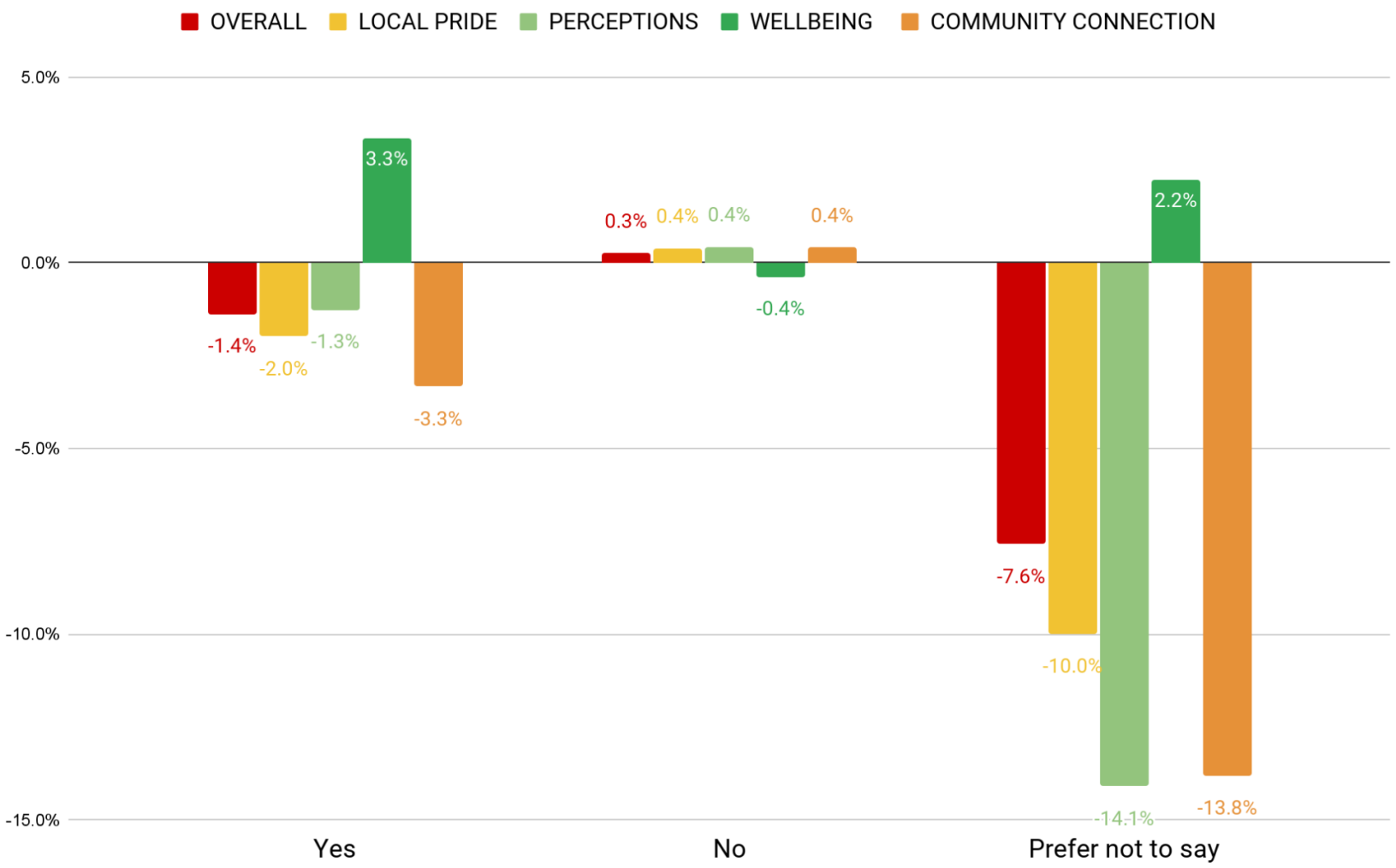


Figure 13. Aggregated Disability Data



Location and Longitudinal Study

As detailed, one of the principal opportunities identified for further exploration is what aggregating data can tell us about locations. If this work were to be scaled up, with data input from other mega events, we would be able to better evidence the long term impact of place based programming. For example, comparing audience data from multiple events happening over time in one location vs multiple events happening over time in another. If we could evidence that one location had received more events-based investment than another, and this event saw stronger outcomes in audience sentiment, then it would go some way in displaying the long term impact of ongoing funding for events.

In order for this comparison to take place, full postcode, or at least specific location data about each audience member must be held. We must also know the location of the work that they were experiencing. When this is combined by externally obtained data points such as events investment levels for that location, we would be able to go some way in proving the power of long term and ongoing investment in events. Aggregating facilitates this analysis as, by taking data from multiple events/sources across a city, you go beyond comparison of events in two different places, and begin to compare general audience sentiment from multiple events between two locations. Further opportunity also lies in the ability to segment these audiences into locals and visitors of the respective cities. To understand, for example, if events taking place in a certain location are more significantly impacting tourism and visitation to that location, or the people who live there year-round.

Moreover, aggregating data also holds some opportunity for longitudinal study of some kind. For example, we know that the three events discussed in this report all took place in a similar geographical location. If we could also prove that they took place in cannon - for example Coventry beginning first, then Birmingham 2022 Festival, then culminating to PoliNations in Birmingham City Centre - aggregated comparison of the datasets could show steady rise in particular outcomes measured through audience sentiment.

This could take two approaches. One: if we knew that the audience member in question had attended the events prior, we would be able to attribute change more directly. Two: if this is not known, a general rise in the outcome across audiences may be enough to prove that multiple events happening in a single location affects general wellbeing, even if we don’t know that audience members in question attended several events.

As noted in this context, all three events undertook in person data collection with audiences as/after they experienced events. In addition COV conducted primary research with a general sample of the population in the city at three points in time, over the course of the event’s lifespan. This latter method provides an easier form of understanding longitudinal effects on a location, as opposed to aggregating data from multiple events happening in one place, over longer periods of time. It is, however, more

costly, and event delivery organisations are less inclined to commission this type of research. This is because it also involves understanding the views of those who did not attend the events they had produced, and is therefore less likely to produce more decisive/positive results.

> **Recommendation:** A middleground for the above dilemma would be to encourage organisations to include questions about whether audience members had attended any large scale events happening in that location in the years prior, with a specific option to identify which events they had attended. This would mean - when aggregated - we would be able to compare the effect of people attending multiple events, to those attending events one-off, across multiple data sources. This would in turn prove long term engagement as a valuable endeavour. This necessitates events to buy in to the concepts and benefits of aggregation.

Event Type Comparison

Another clear opportunity afforded by events aggregation would be for comparison across event types. The delivery models of events discussed as part of this project often involved devolution from the central organisation to subcontractors who undertook delivery of certain 'sub'-events/projects/comissons. In all three of the audience surveys delivered by the events, the audience survey return was linked to the 'sub'-event at which it was collected. If external information was acquired about the type of engagement undertaken at the 'sub'-event - participatory vs non; breakdown of artform/s; free vs ticketed; outdoor vs venue; sport vs culture - then we would also be able to better understand the differing impacts and differing audience across these sub-events. This would then enable programmes in the future to make informed decisions about the type of events they should look to programme to engage specific audience groups.

As has been highlighted at several points the opportunity aggregation provides in this context is the ability to use data from multiple, disparate sources to prove a central insight. Analysis of this kind would allow us to test hypotheses within the events sector to inform future programming, for example:

- Free events better engage those from low socioeconomic backgrounds
- Outdoor events are more inclusive than those which take place in venues
- Sports events are more popular among men, and cultural events among women
- Age affects preference in artform

In order for this to take place we would need to be able to attribute some of the above categories to the events listed on the audience survey returns. In the B22 data, for example, data on the artform of each of the events is available via the activity reporting forms completed by each of the subcontracted organisations. This information is present in the data shared as part of this project and with further work the audience returns from this event could be categorised and compared according to the art

forms they engaged with. These data points were absent from the data shared COV and UBX but could be uncovered via further data sharing or may be present in reporting outputs.

Understanding Data Collection Bias

By aggregating events based audience data collected from multiple sources, and collected in multiple ways, we are also able to ascertain the impact of data collection bias across strands of data collection. If we are able to categorise the method of data collection for each audience return, then we would be able to understand how, for example, digital data collection bias certain demographic groups, or how samples differ if they were collected by professional fieldworks vs if they were collected by volunteers from an organisation, or if audiences filled them in independently on site. This would in turn further help leverage funding for evaluators to deploy professional fieldworks, as this ultimately delivers the most accurate results.

Intersectionality

Intersectionality most commonly refers to the various and overlapping ways in which people can be marginalised in society. Often, as a result of experiencing multiple forms of marginalisation, these are the individuals who stand most to gain from events based interventions of the kind we have discussed. However, as a result of the sample sizes often collected as part of single events, confidence levels are often not high enough within intersectional groups to be able to draw conclusions or provide evidence supporting the narrative of their experience/s. However, the more data aggregated, the more possible it becomes to look into intersectional experiences, and understand how their experiences differ from those of the wider sample.

This particularly came to light in the analysis conducted with the data aggregated as part of this project. Audiences who often do not disclose their demographic information were significantly more likely to experience much lower outcome or sentiment scores. Again, sample sizes for ‘prefer not to say’ demographic groups are often too small when taken at single events alone. However, when we aggregate data their experience comes to the surface, as highlighted in the above figures 9 - 13.

Appendix 1: Research Questions

In August 2023 the following research questions were articulated:

1. How accessible is data from major events?
    a. Is it clear who owns the data?
    b. Is it clear who needs to be contacted to get access to the data?
    c. Who can and cannot access the data?
2. What type of data exists?
    a. Raw, data tables, both, neither
3. What are some of the challenges/limitations of individual event data sharing agreements?
    a. What were the consent clauses used by different events?
        i. Are consent processes uniform across the major events?
        ii. Are there specific limitations in using the data due to consent clauses?
    b. Are there any other limitations in terms of using the research insights?
4. How easy is it to align and use the data?
    a. Is data usable?
    b. Is there consistency between datasets?
    c. How easy is it to align data?
    d. What is the impact of editing data for alignment?
    e. Can collective impact be assessed through alignment?
5. Opportunities and challenges of events data - teams come together and disbanded
    a. Uniqueness of events data - what makes events data sets special
    b. Mega events are moments in time not even annual events - finite
    c. Specificity of events data versus venue or theatre data - breath of data, unique to events data
6. Standardised data
    a. Is there a point of getting the data to be similar, or is it okay to be different (ties in with 4e)

Appendix 2: Analysis Master

Appendix 3: Mega Events in 2022 - Aggregated Data

Appendix 4: Mega Events in 2022 - Data Tables