South African Cultural Observatory

Assessment of the DAC CCI Mapping Study
Methodology

Submitted to the Department of Arts and Culture

arts & culture

Department:
Arts and Culture
REPUBLIC OF SOUTH AFRICA

Summary: Assessment of the DAC CCI Mapping Study Methodology

As Cultural Observatory researchers have worked through the DAC (2014) Mapping Study methodology and raw data, a number of issues have emerged. Some of these are things that relate to research design: They are not necessarily wrong, but need careful thought before another study is run (Part A below). Others are related to the way the data was captured and analysed, and do call into question the reliability of the results (Part B below). This last (incorrect data coding and classification into the UNESCO categories) is what makes the study less reliable. It should be noted that errors were not evident in the report itself, but were only discovered on closer examination of the raw data files provided by Plus94.

A. Research Design and Methodology

This section discusses the Research Design and Methodology used by Plus94. All research must make choices regarding design and methodology – there is no one "right" choice of method. However, the choice of method does have consequences for the trustworthiness and accuracy of results. Before (i) using the results in public documents; and (ii) commissioning a follow-up study, the methods used need to be carefully considered.

1. **Method:** The South African study compiled a database of 24 756 CCIs, and interviewed a sample of 2 477 of them. As a national Cultural Observatory we need to use only raw (non-summarised) data of highest standards.

Point to consider: Is it the case that StatsSA does not produce industry-level data relating to the CCIs? (For example, the Quarterly Labour Force Survey does identify employment by occupation type that could be used to track CCI employment much more accurately.) Using national-level, verified data to map the CCIs might be a more reliable approach than the one followed by the 2014 Mapping Study. To our knowledge, no other national-level CCI mapping study has used the telephone interview approach. The advantages and disadvantages of the method are discussed in the full report (South African Cultural and Creative Industries Mapping Study: Review of methods and the way forward), but some disadvantages include:

- Only about 10% of CCI firms contacted by Plus94 were willing to participate, and not all these were willing to answer all the questions (DAC National Mapping Study, 2014:22).
- Low response rates are a problem in any survey research method, but can be especially problematic if a particular group is under-represented. For example, the Plus94 report notes that it was "very difficult" to get non-South Africans to participate in the research.
- Another indication that the sample was somewhat biased is the high average age of CCIs in the sample (13.61 years), despite the fact that (i) most CCIs internationally, and in South Africa, are small firms; and (ii) small, medium and micro businesses in general are known to have a high failure rate in the first five years of operation, thus resulting in a young average age of firms in such industries.

2. **Definitions and Scope:** An important part of any mapping study is to define the CCIs. There are a number of ways of doing this, all of which are currently debated. While the UNESCO Framework for Cultural Statistic (UNESCO, 2009), used in the 2014 study, provides a broad overview it is seldom implemented directly, often because of the lack of detailed statistical data within various national statistical systems.

Points to consider: The UNESCO Framework, while providing general guidelines, explicitly leaves room for individual countries to make their own decisions about what should be included. For example, no other mapping study (or Framework for Cultural Statistics) found included Natural Heritage (game farms/ranches/reserves) in Domain A (as the 2014 study did). Some studies also excluded parts of those Domains that they regarded as having a low concentration of "creativity" (for example, advertising, internet podcasting and video games). These are things that will need to be carefully considered in follow-up studies.

3. Sampling: The Gauteng regional mapping study (2008) on which the 2014 study seems to be based, used stratified (by small versus large firms) random sampling, while it is not clear at all how the sampling was done in the DAC studies. In addition, sampling appears to have taken place while the database was being constructed (from something said at the hand-over meeting with Plus94 in 2015). This means that those firms who were put into the database early on had a higher probability of being chosen for interview than those that were added later.

Points to consider: As indicated in the discussion of Method, sampling bias can reduce the reliability of the study results, particularly for financial data averages (see below).

4. Monthly Turnover and Economic Impact: One of the most important pieces of self-reported information collected was the monthly turnover or income of CCIs in the interview sample, which was used in the calculation of the economic impact of the industry. However, only 1478 (67.5%) of the sample provided this information and the range was huge: from R0 to R30 million. There was also quite a difference in response rates between the various domains (see table 1).

Points to consider: The Economic Impact study was based on average turnover and other financial data provided by relatively few firms with a huge range. Working with such a small data set, with such large variance, can result in less reliable results, especially when the data is further divided into the different domains and sub-groups. This is one of main reasons that we would suggest that a follow-up study considers the use of StatsSA data, with perhaps some more detailed regional studies that focus on the informal sector to complement it, rather than repeating the survey method.

Table 1: Percentage of firms in each Domain who provided monthly turnover data (including those who gave a R0 response)

Domain	Percentage who provided turnover data
Performance & Celebration	68.87%
Visual Arts & Crafts	56.01%
Information, Books & Press	66.41%
Audio-Visual & Interactive Media	73.08%
Design & Creative Services	73.02%

(Source: Reanalysis of DAC 2014 CCI Interview Data)

B. Data capture, coding and analysis

The points discussed above have the impact of reducing the trustworthiness of results, and/or may need to be carefully considered before a follow-up study is run. However, the real problem with the report as it currently stands is the way in which data was captured and analysed. This applies to both datasets that were handed over to the Cultural Observatory:

- The CCI Database, which consists of over 24 thousand entries, presenting an audit of all the CCI firms in South Africa that could be traced. Variables include the firm name, contact details for some, main business activity, UNESCO Domain, Town and Province, and GPS co-ordinates.
- A database of in-depth interviews (using a mixture of telephone and face-to-face methods) conducted with around 2400 firms. This database includes a large number of variables, including information on the demographics of ownership and employees, financial data, markets and income sources, firm types etc.

CCI Database

Misclassifications: In recent re-analysis of the raw data, it was discovered that there are some potentially significant misclassification errors. For example, the Database lists 2051 firms in the Cultural and Natural Heritage domain. *However, 551 of these are artists or art galleries, which should be included under the Visual Arts and Crafts domain, and a further 120 are Libraries, which should be included under the Books and Press domain.* **At least a third (32.7%) of organisations allocated to the Cultural and Natural Heritage domain are therefore misclassified**.

Implications: Detailed analysis of other Domain classifications has not yet been carried out, but the implications of such large-scale misclassification of organisations into the domains are: (i) that the relative sizes of the CCI domains in South Africa may be considerably different from those reported by Plus94; and (ii) that the economic impact of the CCIs (which it is assumed used interview data averages, weighted by the relative size of each domain) is also likely to be different from what is reported.

GPS co-ordinates and mapping: The CCI database seems to have used two different methods to capture the latitude and longitude coordinates used for the physical mapping of CCI firm locations. One method uses "degree, minutes, seconds" and looks like this: 26°12'37.4"S. The other is the decimal method and looks like this: -29.792112. The database contains a mixture of the two methods.

Implications: While not an expert in GPS mapping, it is noted that the map results available online show some South African CCI firms as appearing outside of the country boarders. This may be because of incorrect co-ordinates or because the use of two different methods of expressing the co-ordinates have caused some errors. It would be problematic to release the map as it current appears, however (see image below).

CCIs that appear to be outside of South Africa in the Plus94 Mapping Portal
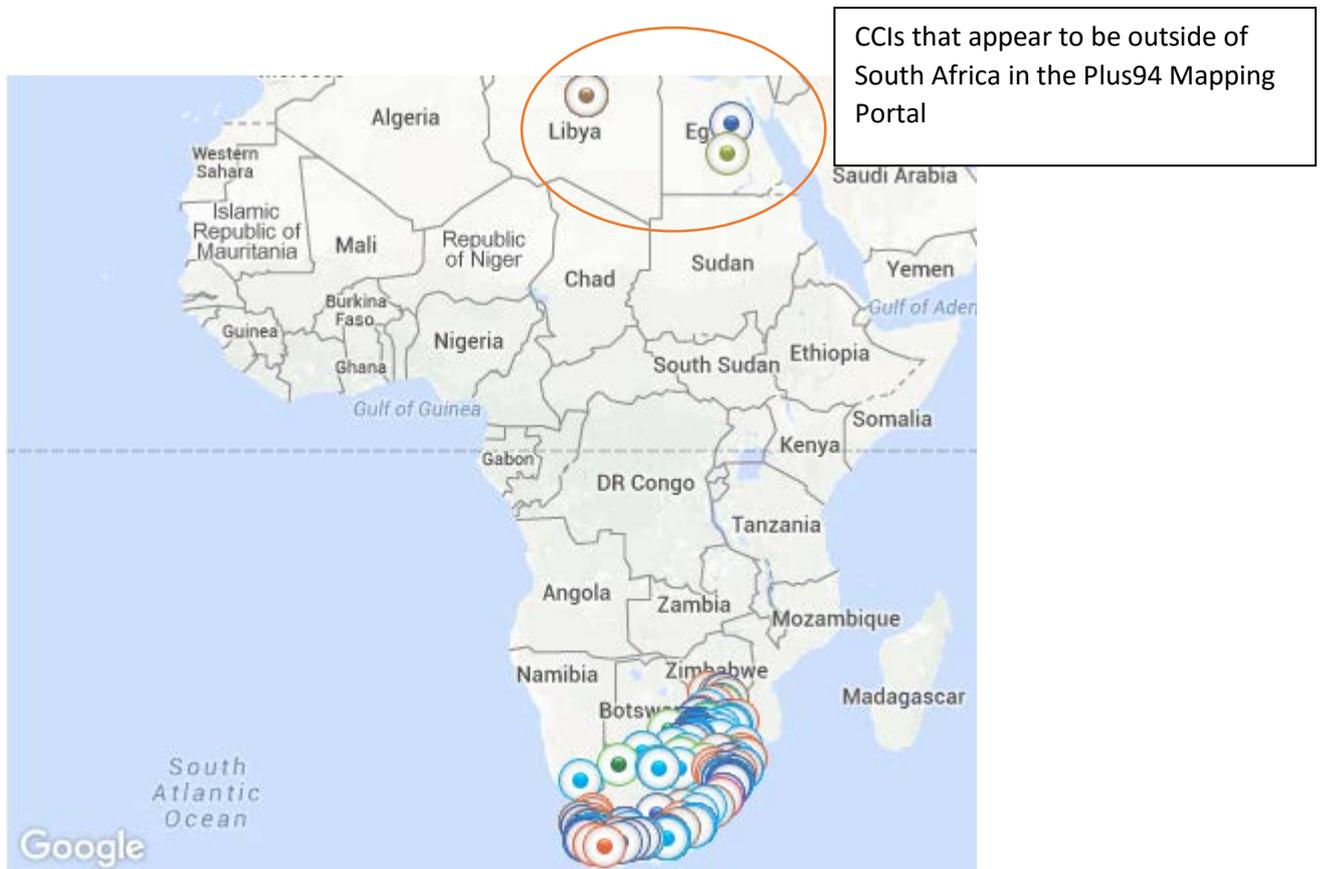
Interview Data

Data capture: The way in which the interview data was captured makes some variables difficult to analyse and requires quite significant re-coding. For example, for Yes/No variables, No is sometimes indicated with 1 and sometimes a 2. In terms of the variable for towns, a "city" listed in Kwa-Zulu Natal province is "Berea". However, this is actually a suburb (There are at least two suburbs called Berea: one in Durban and one in Johannesburg). There are numerous spelling mistakes that make the data difficult to sort for analysis and checking. Some of numbers in the file are recognized by Excel as text, making calculations problematic without re-coding. As with the CCI database, there is some misclassification of firms into the various domains (however, this is not as severe as in the Database).

More worryingly, there are also suspected decimal point errors with the financial data of a small number of firms. For example, taking into account the other financial data provided by the firm, and the number of people it employs, and (where possible to trace) its internet profile, it seems unrealistic that the monthly turnover of even quite a large firm is, for example, is as high as R100 million. It is much more likely that it is R1m. These errors may have been the result of file conversions (from SPSS used by Plus94 to Excel used by SACO) and the use of multiple programmes to capture the data (different for face-to-face versus telephone interviews, according to the Plus94 research manager).

Implications: While the interview data provides a useful starting point, much work is having to be done to re-code it and, in some cases, do a line by line inspection.

C. Concluding Remarks

All research requires methodological choices. Some of the choices made in the Plus94 study may reduce the reliability of the survey in representing the CCIs in South Africa in an unbiased way.

However, a careful evaluation of the methods used is part of developing follow-up studies and improving the next round of research. The issues presented in Part A do not thus invalidate the research.

More serious are the errors made in data capture and analysis, discussed in Part B. This is especially the case with the misclassifications in the CCI Database, that are clearly wrong if the UNESCO Framework is being followed (which is what is stated), and could have big implications for the results. Having said that, the study can be used in some ways, such as:

- The CCI Database needs to be re-coded according the UNESCO categories, GPS co-ordinates fixed, spelling mistakes corrected, repetitions removed etc. However, it does provide a starting point for the construction of a CCI database for South Africa and can be released (after correction) with a request to those CCIs who do not appear on the database yet to add themselves.

- After significant cleaning and re-coding, some of the interview data can be used for research relating to the nature of this sample of South African CCIs (number of people employed, company type, education of employees, ownership patterns etc.). Further analysis can be used to tell us something about these firms, although the lack of data reliability in some variables will have to be noted.

In our opinion, the CCI Database needs to be reclassified as soon as possible, with all the firms put into their correct domains, so that we can assess what difference this makes to the overall picture of CCIs in South Africa and, by implication, the Economic Impact and employment results of the study. It is recommended that Plus94 should be involved in this review so that they can explain their original classifications and the thinking behind them.

In taking the Mapping project forward, we suggest that national level data from StatsSA be explored, rather than the survey method used in the Plus94 study. While this might take longer in terms of generating results, it is much less costly, much more reliable, and in line with international best practice.

The alternative is to undertake the above approach **in tandem** with a revised survey-based set of mapping studies carried out in selected local and/or regional (sub-national) areas. This could involve other government departments. There would seem, as discussed, be a real possibility of undertaking such an exercise in the Eastern Cape.

Professor Jen Snowball

Professor Igor Litvine

Professor Richard Haines

APPROVALS FOR THE SOUTH AFRICAN CULTURAL OBSERVATORY ASSESSMENT OF THE DAC CCI MAPPING STUDY METHODOLOGY				
	NAME	**TITLE**	**SIGNATURE**	**DATE**
Prepared and recommended by:	Prof. Richard Haines	CEO		31/03/2016
Submitted by:	Mphikeleli Mnguni	Research Officer		
Recommended by:	Charles Mabaso	Chief Director		
Approved by:	Monica Newton	Deputy Director General		