

## **Standards for Ethnicity Data**

### **Response from the Employment Lawyers Association**

**30 August 2022**

## Standards for Ethnicity Data

### Response from the Employment Lawyers Association

#### INTRODUCTION

1. The Employment Lawyers Association (“**ELA**”) is an unaffiliated and non-political group of specialists in the field of employment law. We are made up of about 6,000 lawyers who practice in the field of employment law. We include those who represent Claimants and Respondents/Defendants in the Courts and Employment Tribunals and who advise both employees and employers. ELA’s role is not to comment on the political merits or otherwise of proposed legislation or calls for evidence. We make observations from a legal standpoint. ELA’s Legislative and Policy Committee is made up of both Barristers and Solicitors who meet regularly for a number of purposes, including to consider and respond to proposed new legislation and regulation or calls for evidence.
2. A Working Party, co-chaired by Shubha Banerjee and Louise Skinner was set up by the Legislative and Policy Committee of ELA to respond to Standards for Ethnicity Data. Members of the Working Party are listed at the end of this paper.
3. References in this paper to the views of ELA are intended to be inclusive of the views of the minority as well as the majority of ELA members. Whilst not exhaustive of every possible viewpoint of every ELA member on the matters dealt with in this paper, the members of the Working Party have striven to reflect in a proportionate manner the diverse views of the ELA membership.

#### **QUESTION 1: “WHETHER YOU THINK THE DRAFT STANDARDS COVER THE MAIN ASPECTS OF DATA QUALITY?”**

#### **GENERAL COMMENTS**

4. In principle, we consider that introducing standards would give confidence to and encourage employers and other organisations to undertake data collection of this nature. This could provide many benefits in an employment law context, both in terms of enabling an employer to understand its ethnicity (and other) diversity data and make appropriate changes, and also in providing evidence proving or disproving disparities that would assist in relation to the resolution of discrimination litigation. However, in ELA’s view, the current version of the standards for ethnicity data (“the **Standards**”) is very much tailored towards individuals who regularly work in the field of statistical analysis. As drafted, the Standards are unlikely to be user-friendly and easily followed by the majority of those who would seek to rely on them, both in the public and private sectors, including HR functions within Government departments, and as such will be of limited application in their current form. Opinion received from private sector

employers supports this view. Further, it is also possible that the Standards may be used by employees and/or employee representative bodies e.g., in a litigation context in order to evidence ethnicity bias or to help formulate recruitment strategies in collaboration with employers etc. Given this, in ELA's view, it is of particular importance for the Standards to be formulated so that they are also accessible to a lay person.

It is also submitted that the Standards as drafted potentially perpetuate misrepresentations of ethnicity, are not user-friendly, are difficult to navigate – particularly given the additional information links embedded within the Standards – and that the Standards lack a clear, logical structure. With that said, we agree with and support the introduction of standards for the collection and use of ethnicity data (and potentially data related to other protected characteristics). However, we believe these should be more readily accessible to users and include embedded context and examples and should avoid perpetuating unhelpful and outdated concepts.

5. It is stated that the Standards apply to Government departments or public bodies who are collecting ethnicity data, analysing differences between ethnic groups and/or publishing ethnicity data. The Standards are also suggested to be useful to those *“outside of the public sector who collect or use ethnicity data”*. We consider that many employers would want to look to the Standards when collating and reporting on ethnicity data, especially in the absence of guidance for ethnicity reporting. Our members have identified an increase in the collection and processing of personal data revealing racial or ethnic origin, religious or philosophical beliefs and data concerning a person's sexual orientation for the purposes of equal opportunities initiatives and diversity monitoring. Sometimes this is in the specific context of an employer's legal or regulatory obligations (e.g., in the legal industry, employers are required to publish diversity data and under new Listing Rules certain in-scope companies will be required to publish numerical data regarding the ethnic diversity of their board and executive management) and at other times in the more general context of ensuring diversity in recruitment, planning, training and conference opportunities, pay and promotions. There is also growing evidence that the Equality and Human Rights Commission (**“EHRC”**) expects employers to encourage employees to provide diversity data in order to enable more effective equality monitoring (see, for example, the [recent legal agreement between EHRC and Jaguar Land Rover](#) which contains such a commitment). In their current form, the Standards may be of limited use to employers for these purposes as many employers will find the statistical language and absence of clear, practical, examples challenging.
6. When sampling clients' views on the introduction of ethnicity data standards, they have expressed overwhelming support in favour of a template or precedent for collecting ethnicity data or, in the absence of this, at least a starting point to be considered by our employer clients when thinking about how to collect and report on ethnicity data. That said, there has been a divergence of opinion on how standardisation of ethnicity data can be achieved and even whether standardisation is the preferred approach (with the alternative being that organisations adopt classifications of ethnicity pertaining directly to their

organisation). Some private sector clients have pointed out that, if the Standards are to be used to support future ethnicity pay reporting, a unified approach to the collection and reporting of ethnicity data will provide the greatest benefit and allow for better benchmarking across sectors, industries and locations.

7. We consider that data privacy in this context is extremely important, as is the applicable legal framework protecting individuals from discrimination. The Standards should make it clear from the outset that users need to take account of UK GDPR and discrimination issues when applying the Standards (and also provide appropriate guidance for users). Ethnicity data collection is a complex, multi-dimensional matter that goes beyond simple data gathering. We also consider it should be made clear that the Standards are of application in the UK only, given the many different approaches taken to the collection of diversity data in other jurisdictions.
8. The Standards contain boxes setting out supporting evidence and guidance throughout, but it is not clear from the Standards that in order for them to be understood in context, especially for those users who are not used to processing data for statistical purposes, such supporting evidence and guidance needs to be read. Given both the number and length of supporting evidence and guidance notes, we anticipate that employers are unlikely to read these. We would recommend that the need to read the supporting evidence and guidance to aid understanding of the Standards is made clearer, but in any event, summaries of the key points from the supporting evidence and guidance should be inserted within the relevant sections of the Standards. Private sector clients have commented that if information is necessary and relevant to the Standards, it should be contained within the Standards themselves so as to make the Standards easy to use. However, this needs to be balanced against the overall length of the Standards, which in its current form (and given the lack of logical structure) was felt by private sector clients to be approaching an excessive length.
9. We note that the Commission on Race and Ethnic Disparities which suggested the introduction of the Standards also recommended that “*every level of disaggregation adds analytical value providing that it remains possible to draw meaningful comparisons*”.<sup>1</sup> Disaggregation of ethnicity data is further championed by the Race Disparity Unit (“**RDU**”) and is the RDU’s preferred approach, however this is not reflected in the Standards or even referenced in any detail, which we believe would be helpful to users.
10. The Department for Business, Energy & Industrial Strategy (“**BEIS**”) is due to publish guidance for employers on voluntary ethnicity pay reporting. Given the potential overlap with this consultation, we would recommend that the BEIS guidance is consistent with or even adopts the Standards (provided the comments in this response are taken into account) to ensure employers have a good framework of guidance for processing ethnicity data. ELA submitted a

---

<sup>1</sup> The Commission on Race and Ethnic Disparities’ independent report, Foreword, introduction and full recommendations updated 28 April 2021.

response to the consultation undertaken by BEIS on this issue in January 2019 which sets out several comments and recommendations which are equally pertinent to the content and design of the Standards, and which can be found [here](#) (in particular, paragraphs 2.5, 6, 7.1 – 7.2, 8.2 and 9.2 – 9.6).

11. As highlighted above, reference to compliance with data protection principles is notably absent and reference to compliance with the UK GDPR is vague in the Standards. This has been raised as a point of particular concern for both public and private sector employers. ELA considers that it would be helpful to include an explanation at the beginning of the Standards about UK GDPR and how this legislation applies in relation to the collection of ethnicity data. Data protection concepts should be explained such as: personal data, special category personal data (which ethnicity data is considered to be), the lawful bases for processing, the requirements to be able to process special category personal data and the data protection principles users should be considering. This should be drafted in conjunction with the ICO to ensure alignment. Guidance or examples of where employers may seek to rely on the substantial public interest basis for processing ethnicity data under paragraphs 8 and 9 of part 2 of Schedule 1 of the Data Protection Act 2018 and guidance on whether users will be required to undertake a data protection impact assessment would be useful. The ICO's Employment Practices Code states at section 2.5.2 that data processed for equal opportunities monitoring should, where practicable, be kept in an anonymised form. The ICO's Employment Practices Code does not align with a growing trend for employers to measure progress (or lack of) across participation and opportunities at work. The practice of measuring progress is, we believe, becoming an industry standard of good practice for inclusion and diversity programmes, particularly for employers with a global presence. The Standards could be an opportunity to provide clarity on when processing identifiable data could be justified.

## **INTRODUCTION**

### **The Importance of ethnicity data**

- 11.1. This section states that “the amount of ethnicity data and the quality of ethnicity data is important”. We would recommend that prior to this statement there is an acknowledgement of why ethnicity data is important in the first place (the last sentence refers to the need to reduce disparities between ethnic groups but we would suggest that this is a narrow view of why ethnicity data is important). The Government might consider that this falls outside the scope of the Standards (particularly if collection is mandatory for a particular organisation/sector). However, it is important that any user of the Standards, including employers, first addresses to the issue of why they are collecting ethnicity data i.e., to assess participation and recruitment practices, improve representation, produce fairer outcomes – not just between ethnic groups but across all ethnic minority groups.
- 11.2. Furthermore, in order to comply with data protection principles set out within the UK GDPR, such as data minimisation and purpose limitation,

it is important to understand why ethnicity data is being collected to ensure that the users of the Standards are only collecting ethnicity data for the specified purpose(s) and limiting the data collected to what is reasonably necessary to achieve that purpose. As the Standards do not currently acknowledge the need to consider this at the outset, there is a risk that the Standards simply promote data collection for data collection's sake, which may, in turn, lead to lower engagement, poor disclosure rates, inaccurate analysis/reporting and potentially non-compliance with data protection principles where users may not necessarily understand the need to comply with these.

- 11.3. If such information regarding purpose is included in the Standards, it will allow employers to use the same language in their communications with employees and better provide comfort and certainty to employees on the purpose of the data collection. The purpose and value of the data collection is of paramount importance. This point has been raised by a number of private sector clients. Further, feedback from clients suggests that where the purpose and reasons behind the collection of data are fully explained, individuals are more likely to provide a response.

#### **Who the standards are for**

- 11.4. We would recommend including specific reference to the other types of users outside of Government departments or public bodies who may use the Standards, such as employers, if the Standards address the concerns raised within this response. It may not be abundantly clear to employers on reading the Standards that they are able to voluntarily adopt the Standards, which may result in a lower take up of the Standards outside of Government departments.
- 11.5. We would recommend that the Standards specifically address within this section whether or not the Standards are mandatory when a Government department or public body outsources activities to a third party. Our view is that a reasonable stance may be that the Standards should apply when application of the Standards would otherwise have been mandatory, but for the process having been outsourced.

#### **Using the Standards**

- 11.6. The first paragraph suggests that ethnicity data of itself is a sensitive topic. We would suggest that this paragraph should contain an acknowledgement that any perceived or actual sensitivity may arise from the reason for collecting ethnicity data (e.g., has this been properly communicated, is the data relevant/required, how will the data be used etc.) as opposed to the collection of such data in and of itself. This is an important point to highlight for users of the Standards as it will more accurately represent any perceived or actual sensitivity

around ethnicity and will contribute to better understanding and more confident use of the Standards.

- 11.7. We suggest that the second paragraph on drawing conclusions from unreliable ethnicity data due to small numbers of participants requires further explanation of how sample size impacts upon the usefulness of the data and the ability to draw conclusions, as explained in the supplementary guidance.<sup>2</sup>

### **Monitoring the use and impact of the Standards**

- 11.8. We note that for other users of the Standards where the use of the Standards is not mandatory, it is not clear exactly what would be required when it comes to monitoring the use and impact of the Standards (e.g., would non-mandatory users be required to inform the Equality Hub/Office for Statistics Regulation (“OSR”) when they collect ethnicity data, would they need to confirm compliance with the Standards and would failure to comply with OSR guidance have any wider/financial impact). Please refer to the response to question 2 below for further comments on the monitoring of ethnicity data statistics.

### **Key Considerations: Quality**

- 11.9. Whilst directed at Government departments and public bodies, it is likely that employers will look to these Standards when processing ethnicity data in their organisations. For some employers, there will be a legal or regulatory obligation to collate such data. However, there will be many employers who are not subject to any specific obligation to collect ethnicity data. Users of the Standards will need to ensure their processing of ethnicity data complies with data protection law, with the primary question being whether users have a lawful basis to process the data. We therefore recommend changing the order of the current sections 2 and 3 of the Standards so that Section 3: Trustworthiness comes before Section 2: Quality. This would also align with a recommended approach to processing ethnicity data which requires consideration of: (i) why is the data needed / can you lawfully collate such data? (ii) how do you ensure that the data you collate is quality data, and (iii) how do you analyse, and (if appropriate) report that data? As above, we submit that this would also allow for employers to better communicate to employees the purpose of their data collection and the intended outcome, as well as what such data will not be used for. A number of private sector clients commented that a “prefer not to say” option has had one of the highest rates of response when collecting such data and that clarity around the purpose and process might reduce the number of “prefer not to say” responses.

---

<sup>2</sup> The RDU’s Research and analysis “*Ethnicity data: how similar or different are aggregated ethnic groups?*” published 22 December 2020.

## **Data collection**

### **Collecting ethnicity data should be a priority**

- 11.10. There is reference to the collation, analysis and reporting of ethnicity data being a “priority”. We consider that this statement could be misleading as it suggests that users of the Standards should process ethnicity data as much as possible. We would recommend adding a statement to remind users of the Standards, that before they can begin processing ethnicity data, they need to ensure that they have complied with the relevant data protection laws and the principles to ensure that users have a lawful basis for processing ethnicity data. The Standards should also make reference to the risk of identification of an individual where the response groups fall below a particular number, and state that employers should exercise caution where this might be the case and possibly consider collecting aggregated data instead.

### **At the start, think about how you will use the ethnicity data you collect**

- 11.11. We would recommend that the heading “At the start, think about how you will use the ethnicity data you collect” could be amended to “At the start, think about the purpose(s) for which you will use the ethnicity data that you collect”. We have recommended this subtle change in wording as it uses terminology found in UK GDPR such as “purpose”. This focuses minds and will hopefully encourage consideration of the data protection principles.
- 11.12. The question “How robust do you need the results to be?” could be amended to read “What are the reason(s) for the data collection?”. It is important for users of the Standards to address their minds to this from the outset, as recommended by the purpose limitation principle under UK GDPR. The reason the ethnicity data is collected is also integral to the quality of the data collected and ensuring that the quantity of any personal data collected is kept only to what is reasonably necessary for the reason for which it is collected. Further, the phrase “How robust do you need the results to be” is unlikely to be an appropriate question. If, as stated, the guiding principles of the Standards are “quality, trustworthiness and value”, all users (including employers) should be focused on collecting the most robust (i.e., high quality, trustworthy) data as possible.

### **Collect ethnicity data using the Government Statistical Service (“GSS”) harmonised standards, or more detailed groups that you can align with the harmonised standards**

- 11.13. The Standards recommend collecting “ethnicity data using the GSS harmonised standards, or more detailed groups that you can align with the harmonised standards”. It would be useful for users of the



Standards to have a link or a copy of the GSS harmonised standards within the Standards so that they are easily accessible. Our recommended preference is for clear, practical examples in this section (and throughout the Standards) which would aid users of the Standards such as employers in understanding and implementing the Standards, rather than just including links to supporting evidence and guidance.

- 11.14. The Standards recommend providing an “option for respondents to write in their ethnicity”. We consider that further explanation could be provided here, for example, the Office for National Statistics (“**ONS**”) have confirmed that in relation to the 2021 census, ‘search as you type’ functions within an option box helps remove errors that can occur when there is free text typing available (as these errors make it difficult to compare data).<sup>3</sup> ONS, however, also highlight that ‘search as you type’ text boxes may impact responses if the option that applies to the responder does not appear. One recommendation may be to use a mixture of tick boxes and then a free text option box for anything not listed. These methods allow for the disaggregation of ethnicity data which is the recommended approach. This kind of useful practical guidance that users can take into account is currently missing from the Standards and would be welcomed, particularly as clients have already expressed concern about anything that will make the data less readily usable or inappropriate for comparison.
- 11.15. It is not clear from this section, without having read the GSS harmonised standards, that there are different standards for different parts of the UK and the reasons why these standards differ. We suggest that this should at least be signposted to users of the Standards so that they are aware that their approach may need to change depending on where in the UK they are collecting ethnicity data.

#### Collect data on religion and national identity

- 11.16. There is a suggestion within the Standards to collect data on national identity and religion when obtaining ethnicity data. Whilst we understand the rationale of allowing people to express their full cultural identity, we consider that asking for this information as part of understanding an individual’s ethnicity data, is unusual and may be unhelpful for the purposes of data analysis if to do so results in too much differentiation between the data that is obtained. See also the point we make in the paragraph below about data minimisation. The same recommendations about including ‘search as you type’ boxes apply to the collection of national identity and religion data, if individuals are given the opportunity to write in their own religion or national identity, as set out above for ethnicity data. Certain clients had

---

<sup>3</sup> ONS data and analysis from Census 2021, National identity, ethnic group, language and religion question development for Census 2021.

expressed an intention to consider the intersectionality of their employees and to also break down data by geographical location, but this was seen as something that they would only be able to do once they had successfully implemented systems to collect and analyse ethnicity data itself. It was therefore felt that to look at other characteristics would over-complicate the Standards at this stage.

- 11.17. We consider it appropriate to include references to data protection principles within this section, given that, as currently drafted, the document may encourage users of the Standards to collect more personal data than may be necessary through phrases such as “consider collecting data on national identity and religion. This improves the acceptability of the ethnicity question”. If users of the Standards are only collecting data on national identity and religion to ensure answers are provided to questions on ethnicity data, this is not compliant with UK GDPR and the principle of data minimisation. Users should only collect personal data which is reasonably necessary to achieve the intended purpose(s), which in this case is likely to be to assess participation and recruitment of ethnic minorities, improve representation of minorities, and produce fairer outcomes. Focus should instead be placed on considering how users can frame questions on ethnicity data, for example providing supplementary background information about the reason for data collection and how the data will be processed and reported to encourage disclosure of information, against questions designed to elicit ethnicity data – this aligns with the aim for the Standards of “*trustworthiness*”. This approach should be repeated where this statement appears throughout the Standards.

Ask people to report their own ethnicity, where possible

- 11.18. We do not consider that “third-party” or “proxy” ethnicity reporting is effective. It is not appropriate, for example, to “use visual appearance” or “use an algorithm based on name and location” to provide someone’s ethnicity. Furthermore, the Standards should signpost that the use of algorithms may be classed as automated decision making or profiling under UK GDPR, which has increased protection due to the impact such processing can have on an individual’s personal data. The adoption of such methods contradicts the fact, as the Standards highlight, that ethnicity is a “self-defined and subjective concept”. Any data collected in this manner would be of low quality and not necessarily reflect the ethnicity of the relevant person – so arguably there is no justification for, or value in, reporting data on a third-party or proxy basis. From an employment law perspective in particular, reporting data on such a basis could be offensive, inaccurate and could

give rise to discrimination claims.<sup>4</sup> This was a concern expressed by the majority of clients asked. We are also of the view that the Standards should more clearly make reference to the need to consider the collection of data alongside employment law issues such as race discrimination.

- 11.19. There is reference to using “best practice when collecting data to increase response rates and reduce the amount of missing ethnicity data”. This is unhelpful on its own and a summary or examples of what best practice is, how response rates may be increased and how missing ethnicity data could be reduced would be helpful if included. In particular, guidance on how to increase the response rate to ethnicity data questionnaires and also on how to reduce the number of “prefer not to say” responses would be of value.

#### Design data collections to increase the representativeness of ethnic groups

- 11.20. It could be useful to change phrases from “you might use best practice” to “we would recommend best practice is used”. This subtle tweak to the language is likely to increase uptake of the use of best practice.
- 11.21. The recommendations regarding the use of sample boosts or different survey techniques such as snowball sampling are unlikely to be familiar/relevant/meaningful for private sector employers and an explanation of these techniques would be useful.

#### Use data linkage to improve ethnicity data quality

- 11.22. In relation to data linkage, the focus should be on the purposes for which the more accurate/complete data was gathered and the permissions sought/provided. For example, if an employer sought ethnicity data for reason X, would it be entitled to use the linked data for reason Y? There is no acknowledgement in the data linkage paragraph of the legal parameters of processing ethnicity data in this way, for example this is not compliant with the purpose limitation principle within UK GDPR, and this should be signposted within this section.

---

<sup>4</sup> There is currently an investigation by the ECHR into Pontins, a holiday park operator in the UK, owned by Britannia Hotels following a whistle blower allegation that Pontins operated an internal “*undesirable guests*” list consisting of mainly Irish names. This was part of an alleged policy to refuse bookings by Gypsies and Travellers.

## **Data analysis**

### **Use harmonised categories for analysing ethnicity data**

- 11.23. We understand from the RDU the preferred approach, wherever possible, is to disaggregate ethnicity data.<sup>5</sup> This section should be updated to reflect that disaggregated ethnicity data is the preferred approach by the RDU including any reasons for this preference. This explanation will encourage users of the Standards to adopt the approach of disaggregating ethnicity data wherever possible and to understand why this approach is preferable to using the 5 aggregated groups.
- 11.24. We would suggest the limitations for all 5 aggregated groups should be listed (not just for Black/Black Caribbean/Black African) to ensure users of the Standards are fully aware of the ramifications of using these groups on the resulting analysis/reporting. This is preferable to linking to the supporting evidence/guidance and will ensure that the Standards are a practical and informative document for users.
- 11.25. Guidance on the Standards' position in relation to binary reporting would also be helpful, as a number of clients felt that they were only comfortable/able to analyse and report in this way currently.

### **Use appropriate comparators in your analysis**

- 11.26. We would suggest that examples are provided to ensure the Standards are a practical and informative document for users of the Standards. Any guidance the RDU can provide on whether there is a particular threshold in response rates which is required for analysed data to be considered accurate or usable is welcomed.

### **Find out whether the geographic clustering of some ethnic groups has produced counterintuitive results**

- 11.27. There is a suggestion that users of the Standards should "consider whether the disproportionate concentration of some ethnic groups in urban areas has led to counterintuitive results". The language used here is unhelpful. What is deemed "disproportionate"? Why is there now reference to "urban" areas? The Standards could repeat the reference in the heading to "geographic clustering of some ethnic groups" instead.
- 11.28. There is also reference to a Simpson's paradox, it would be useful to explain what a Simpson's paradox is and provide an example for users of the Standards who are not familiar with statistical language, such as

---

<sup>5</sup> The Race Disparity Unit ("RDU")'s Research and analysis "*Ethnicity data: how similar or different are aggregated ethnic groups?*" published 22 December 2020.

employers. Alternatively, it may be appropriate to exclude from the Standards such specialist statistical terminology to minimise the risk of confusion and lack of understanding amongst users.

Consider whether you measure differences between ethnic groups by analysis of raw data or after adjustment to take into account other socio-economic and demographic factors, or both

- 11.29. The following statement is broad, unhelpful and unlikely to be informative for users of the Standards: “For example, people in ethnic minority groups tend to be younger than white British people and are more likely to live in large urban areas.” Such broad generalised statements (without any statistical evidence in support) are likely to undermine confidence in the Standards.
- 11.30. Again, statistical references (“regression analysis”, “collinearity”, “confidence intervals” etc.) are unlikely to be familiar/relevant/meaningful for users of the Standards such as employers. We would recommend including explanations and examples.

### **Data reporting**

- 11.31. Please refer to paragraph 11.18 above on proxy reporting/imputed ethnicity records. The Standards state that users should report if there has been any proxy reporting or imputed ethnicity records when presenting the data. We think it is important that we confirm again that we do not recommend of these methods to collect ethnicity data. This is a view shared by clients who were asked to comment on the Standards.

### **Key Considerations: Trustworthiness**

#### **Data collection, reporting and analysis**

Collecting ethnicity data in a respectful way – it should support public interest

- 11.32. We consider that the reference to collecting data “in a respectful way” is at odds with the use of third party/proxy reporting (please refer to paragraph 11.18 above). We do not recommend the use of third party/proxy reporting for ethnicity data.
- 11.33. The Standards refer to the need to regularly review ethnicity data. It would be helpful to see some guidance on what the RDU considers to be a regular interval for such review.

### **Key Considerations: Value**

- 11.34. We do not have any comments on section 4 of the Standards.

## Summary of Ethnicity Data Standards

- 11.35. We do not have any comments on section 5 of the Standards (on the basis that this is a summary of the earlier sections and all substantive comments have been addressed above).

### **QUESTION 2: “WHAT YOU THINK OF OUR PLAN TO MONITOR THE IMPACT AND USE OF THE STANDARDS IN COLLABORATION WITH THE OFFICE FOR STATISTICS REGULATION?”**

12. The OSR is the regulatory arm of the UK Statistics Authority. The UK Statistics Authority was established under the Statistics and Registration Service Act 2007. The UK Statistics Authority is an independent statutory body, that operates at arm’s length from the Government as a non-ministerial department.

13. The OSR’s principal roles are to:

- 13.1. Set the statutory code of practice for official statistics;
- 13.2. Assess compliance with the code;
- 13.3. Award the national statistics designation to official statistics that comply fully with the code; and
- 13.4. Report any concerns on the quality, good practice and comprehensiveness of official statistics.

### **OSR and Government Collaboration**

14. Collection of data regarding ethnicity is a sensitive topic, especially when collectors of ethnicity data are asked to provide details of the use and purpose for which such ethnicity data is being collected. Given the OSR was established to be independent of the Government, any proposed monitoring on the impact and use of the Standards may be more likely to be accepted if the OSR carried out the monitoring itself and not in collaboration with the Government.

### **Benefits of the OSR Monitoring the Impact and use of the Standards without Government Collaboration**

15. There are a number of advantages in the OSR alone monitoring the impact and use of the Standards. These are:

- 15.1. The OSR is independent of the Government;

- 15.2. The OSR has a track record in challenging official statistics;
- 15.3. The OSR's Code of Practice provides guidance on how data should be collected and will help provide benchmarks for the collation of quantitative data;
- 15.4. The OSR has a track record in ensuring that any data collected is not misused;
- 15.5. Members of the public, including private organisations, can report concerns to the OSR;
- 15.6. The voluntary application of the code and engagement with the OSR could assist those in the private sector; and
- 15.7. If there is concern that any data produced will not be accurate, the OSR's involvement could provide confidence.

### **Challenges with the OSR Monitoring the Impacted use of Standards without Government Collaboration**

16. There are several challenges that the OSR would face as the sole body monitoring the impact and use of the Standards. These are:
  - 16.1. Does it have the resources to monitor the impact and use of the Standards bearing in mind its other case work?
  - 16.2. Will the monitoring be focused on official statistics rather than assisting the private sector?
  - 16.3. The reports produced by the OSR could be too macro-level and not assist private sector organisations who will be applying the Standards to their own organisations.
  - 16.4. There are a number of different organisations in the private sector in terms of size, sector focus, and jurisdiction. Only the very large private sector organisations will have the resources to produce data and engage with the OSR. This could lead to little voluntary participation and therefore challenges for the OSR to provide meaningful monitoring of the impact and use of the Standards in the private sector.
  - 16.5. The private sector will need time to understand and apply the Standards and may be put off by scrutiny from the OSR.
  - 16.6. The OSR will monitor the use and impact of the Standards based on quantitative data. However, in trying to understand the data, qualitative data will be required as well.

16.7. There may be less engagement by both individuals from different ethnic groups as well as organisations if there is monitoring by the OSR. A great deal of communication and reassurance will have to be provided as to its role. In particular, private sector organisations will need to understand that the OSR will not intervene in their organisations nor name them publicly as they try and apply the Standards. Nor will there be enforcement action if genuine mistakes are made.

16.8. The Government and the OSR are keen to collect data which can be compared. There is therefore an emphasis on collecting data in a similar way. However, the private sector is full of unique organisations that vary in size and resource. It will be harder for the private sector to collect data which can be compared with the public sector or other organisations in the private sector. It needs to be acknowledged that one size does not fit all.

### **Conclusion**

17. The OSR exists primarily to challenge the trustworthiness, quality and value of official statistics relied upon by the Government. We query if it can expand its purpose to assist with scrutiny of the impact and use of the Standards in the private sector or whether a different body should be responsible for the same. It is important to monitor the impact and use of the Standards in the private sector as well as the public sector, and it is important that whichever body is chosen is independent of the Government and monitors both the public and private sector.



**QUESTION 3: “SHOULD WE DEVELOP SIMILAR STANDARDS ABOUT DATA FOR OTHER PERSONAL CHARACTERISTICS?”**

18. Yes, generally speaking (provided the above comments on the Standards are taken into account when doing so), although for certain characteristics (e.g., age, marriage and civil partnership) it may not be necessary as the delineations are more straight forward and less open to interpretation.
19. Considering the purposes of producing a data standard, the characteristics that most lend themselves to the production of a standard are those where there is potential for inconsistency in data collection, and/or where there is particular sensitivity around asking people about that characteristic (as with ethnicity) such that organisations may benefit from support and advice on how best to collect such data and/or where those who share particular characteristics may suffer from a particular group disadvantage.
20. The reality is that many – if not most – organisations (both public and private sector) are already collecting such data.
21. They are currently doing so for a number of reasons, including, recruitment, service planning, service delivery, workplace equality, compliance with legal obligations (a particular example in the public sector being the requirement to comply with the public sector equality duty for which such data is essential) – and for many purposes, including, internal reporting, external reporting (such as to regulators) and for submission to workplace equality benchmarks such as Athena Swan<sup>6</sup> and the Disability Confident employer scheme.<sup>7</sup>
22. At present, organisations are collecting this data within the bounds of many different, and often overlapping/conflicting sets of regulations and standards, that were either: (i) not designed to govern the collection of often special category data on such a large scale; or (ii) create an express presumption against collecting such data on such a scale.
23. In our view and as referenced above, organisations of all sizes and in all sectors would benefit from a consistent set of standards to follow. This will achieve the dual aims of protecting the security of such data (in accordance with the Data Protection Act 2018 and UK GDPR requirements) whilst minimising the administrative burden on organisations that multiple sets of standards can create, assuming that the standards are flowed through to the private sector.
24. This does however give rise to two additional, potentially far-reaching, questions:
- 24.1. What characteristics require standards; and

---

<sup>6</sup> <https://www.advance-he.ac.uk/equality-charters/athena-swan-charter>

<sup>7</sup> <https://www.gov.uk/government/collections/disability-confident-campaign>

- 24.2. How do you ensure the creation of long-term comparable data in a world where the characteristics with which one identifies may change over time.

## **WHAT CHARACTERISTICS REQUIRE STANDARDS?**

25. The question mentions “personal” characteristics, but what do we mean by that?

26. It seems to us there are two possibilities, we either mean:

26.1. Protected characteristics (within the meaning of the Equality Act 2010); or

26.2. Personal characteristics, beyond the characteristics that are protected under the Equality Act 2010.

### **Protected characteristics**

27. The benefit of focussing on the protected characteristics recognised in Chapter 1 of the Equality Act 2010 (“**Chapter 1**”) is that it provides a stable, recognised list of identifiers around which meaningful and comparable data can be collected.

28. If the intention of the proposed Standards is to enable Government and multiple organisations to collect data that can be compared both over a period of time and to allow comparison between those organisations, it is essential that the foundations upon which such data collection is built are as stable and consistent as possible.

29. Of the protected characteristics in Chapter 1, we consider that data standards would be helpful for protected characteristics, with the exceptions of age, and marriage and civil partnership status. This is because these categories are more straightforward: an individual is either married or not, or in a civil partnership or not; and these categorisations are generally not controversial or difficult -there is much less need then for guidance for employers in how to collect such data.

30. We note, however, that not all of the protected characteristics in Chapter 1 are without issue. In particular, the term “*gender reassignment*” has been criticised as outdated.

### **Personal characteristics**

31. If ‘personal characteristics’ refers to something broader than the protected characteristics in Chapter 1, then what is it that is actually meant?
32. A ‘personal characteristic’ potentially encompasses a vast array of data, from personal appearance to how one’s leisure time and financial resources are spent and beyond, many of which will not be apt for incorporation into a standard.
33. In order to develop a standard about such personal characteristics, Government would need to develop a list of the characteristics about which organisations

collect or may wish to collect data but which are believed to require guidance. On the basis, therefore, that there needs to be a list, what characteristics are to be on that list?

34. In determining the answer to this we need, in our view, to consider the purpose for which the data is being collected in the first place - that will dictate which characteristics an organisation needs to be aware of in order for the data to achieve the purpose. For example, if the purpose of the data collection is to identify where additional educational resources ought to be deployed, it may be important to assess the relative educational attainment and socio-economic status of an area (although, in doing so, it is important to avoid conflating correlation with causation). Governmental and private sector organisations are likely to have a myriad of different purposes in collecting such data, but we take the view that understanding and redressing inequalities is likely to be one of the primary purposes.
35. Whilst the list of characteristics ought to be the subject of further consultation, the 2021 Census<sup>8</sup> in our view provides a useful starting point. It has questions focussed on the characteristics protected by the Equality Act 2010, but it also contains a myriad of other questions which are clearly designed to discern other significant “personal” characteristics of respondents.
36. For example:
- 36.1. Educational status and attainment (see questions 28 - 31);
  - 36.2. Occupation (see questions 33 - 51);
  - 36.3. Residential status and standard of accommodation (see questions H7 - H14); and
  - 36.4. Socio-economic status (see, for example, question 40).
37. The additional challenge that comes with adopting the approach of “personal” as opposed to the protected characteristics in Chapter 1 is the inherent difficulty in selecting the appropriate categories within each characteristic and how those categories relate to the purpose for which the data is being obtained. For example, one could describe parental status as a personal characteristic, but “parenthood” is extremely broad – it can range from informal parenting of a partner’s child, through to parenting your own biological children, through to adopting, through to fostering, through to acting *in loco parentis* for a grandchild. In order to arrive at the appropriate categorisation, we need to be clear about what the purpose of the data is. As the purpose changes so too will the questions asked. This would pose a challenge for a standardised approach to collecting such data.

---

<sup>8</sup> <https://census.gov.uk/>

38. The same is true for characteristics such as socio-economic status, which is so often confused with socio-economic background (the former being the status you currently associate with yourself and the latter being the status in which you grew up). The answer to that question, again, lies in what purpose the data is being collected for. If it is, for example, to determine where best to deploy educational resources then socio-economic background is largely irrelevant, what matters is socio-economic status. Some aspects of social deprivation for example may be difficult to fit within a standard.
39. These issues will need to be considered in detail in relation to any consideration of introducing standards for further data collection.

### **HOW DO YOU ENSURE THE CREATION OF LONG-TERM COMPARABLE DATA?**

40. A degree of consistency – afforded by a strong foundation to the data set – is essential if the data collected is to remain useful over a period of time and across data sets.
41. Using a recognised and stable list of identifiers – such as those characteristics protected by the Equality Act 2010 – provides a recognised and stable foundation, whilst having the added benefit of a wealth of supporting case law which provides a framework to determine where one sits within each of those protected characteristics.
42. The difficulty that arises – and one which we suggest the proposed standards should address – is ensuring consistency in how different individuals assess their position within each protected characteristic. Consistency is more likely to follow where different organisations are able to utilise the same standards and have the benefit of the same specific advice and guidance when undertaking data collection. In the absence of this, differing individuals may make different assessments of what it means, for example, to be disabled – particularly when those differences are multiplied many times over – which will result in skewed and incommensurable data.
43. In some cases, this is relatively straightforward. Age, for example, is a simple factor of how much time has passed since one's birth.
44. For other protected characteristics, this is very difficult. There are competing views, for example, as to whether the correct measure should be biological/birth sex or identified gender. There is a clear need to consider options for gender beyond the traditional gender binary. As this is a sensitive and polarised topic, the production of a data standard to assist organisations in approaching data collection in the right way would be useful.
45. Another example of where standards would be useful is in relation to collection of disability data. Whilst there is a test in section 6 of the Equality Act 2010 against which one may assess whether, as a matter of law, someone is disabled, that

legal analysis may be very different to whether someone considers themselves to be disabled (either because the individual does not consider themselves to be disabled when the law would regard them as so, or because someone may conflate the subjective concept of ill-health with the objective legal concept of disability). One's perception of whether they are disabled, and indeed whether they meet the test of disability under the 2010 Act, may also change over time. The production of a standard in relation to the collection of disability data would therefore be helpful.

46. Using a list of personal characteristics beyond those protected by the Equality Act 2010 simply compounds these difficulties, as personal characteristics can become increasingly subjective – to continue the example of socio-economic status, an individual may consider themselves to have a “working class” socio-economic status because of their “roots” when on an objective analysis they would be considered affluent and “middle-class”. Equally, one can be both “working-class” and wealthy; they are not mutually exclusive. It depends on what is sought to be measured.
47. The standards ultimately adopted will need to make allowances for these inherent difficulties. Taking disability as an example:
  - 47.1. If, for example, the standards require alignment between those categorised as ‘disabled’ in any data collected and the legal definition of ‘disability’, then any standards introduced will need to assist organisations with how to obtain data which allows that assessment to be undertaken.
  - 47.2. If, instead, the standards permit respondents to “self-identify” as “disabled” then there would need to be clear guidance for respondents as to what that means.
48. Any standards put in place for personal or protected characteristics would need to be user-friendly recognising that the people collecting and interpreting the data are not statisticians (especially if the standards are adopted by private employers) so any standards would need to be easily understood and applied by those not familiar with statistics or statistical theory. Many private sector organisations do not employ anyone with a statistics background, and therefore if the standards are to assist such organisations, discussion of statistical modelling techniques would need to be simplified and/or contextualised by examples, as we mention above in relation to the Standards themselves.

**Members of the ELA Working Party**

Shubha	Banerjee		Co-Chair
Louise	Skinner	Morgan Lewis & Bockius	Co-Chair
Ranjit	Dhindsa	Field Fisher	
Kathryn	Edwards	Gibson Dunn	
Lauren	Piesley	Gibson Dunn	
Julia	Wilson	Baker & McKenzie	
Shuabe	Shabudin	Pinsent Masons	
Christine	Young	Herbert Smith Freehills LLP	
Angharad	Schell	University of Warwick	
Mayer Brown group of solicitors			