Office for Statistics Regulation

Misleadingness: Paper 2

April 2021

Contents

Misleadingness paper 2 - Introduction

Why we did it

At the Office for Statistics Regulation, we are often asked if we consider a particular use of statistics to be misleading. These questions can come from members of the public, politicians and organisations and we welcome them, because the interest in whether uses of statistics are misleading or not shows that people care about the appropriate use of statistics.

We always look carefully at these cases and seek to reach a judgement, but in investigating them, we find it is often not clear what is meant by something being "misleading". The word is used to cover a wide range of situations and sometimes it seems as though the judgement we are being asked to make revolves around the merits of the argument that the user is making, rather than the use of statistics in itself.

So over the last year we have been thinking about the idea of misleadingness – what it is, and how we should approach it in the context of our work. We wanted to go beyond merely technical criteria and think about the impact of uses of statistics on audiences. Our first step was to publish a think-piece in May 2020, which we developed with input from Jenny Saul, a philosopher who has written and thought extensively about misleadingness.

Our think-piece explored three approaches to judging misleadingness:

1: Materiality and intention – an approach which focuses on the significance of the statement being made. What were the intentions of the speaker?

2: Audience – an approach which focuses on audience understanding. Were the audience misled about what the statistics were telling them?

3: Case-based – an approach which focuses on particular features of the presentation of statistics. Is the style of presentation unclear and likely to mislead?

We concluded that the most appropriate definition of misleadingness in the context of our work as statistics regulator was:

"We are concerned when, on a question of significant public interest, the way statistics are used is likely to leave a reasonable person believing something which the full statistical evidence would not support."

We also determined that none of the three approaches was likely to be effective on its own. Instead, the think-piece tentatively concluded that a blended approach was likely to work best.

The paper that follows provides an update on our thinking based on conversations we've had and feedback we've received since we published the think-piece.

Who we spoke to

After the initial publication of the misleadingness piece, we received feedback from a number of sources, including:

- Further input from Jenny Saul, the philosopher who worked on the first think-piece
- Outcomes from a seminar held with Jenny Saul, other philosophers she suggested[1], Ofcom and the Advertising Standards Authority
- A meeting with the RSS Data Ethics Committee
- Individual feedback from the chairs of the RSS Data Ethics Committee
- Feedback from a small number of other individuals

What we found

Overall support for the approach in the think-piece

- Overall, people welcomed the think-piece; it was valued as much as a trigger for discussion as for its content.
- One clear outcome was a recognition of the benefits of bringing together statistical, philosophical and regulatory approaches. Several people who provided feedback commented positively on this way of working.
- Having a clear statement of principles is helpful but we also need to recognise an irreducible complexity. Professor Kevin McConway of the Open University pointed out to us that it will always be difficult to produce a definitive document that describes every possible situation of misleadingness.

Distinguishing production and use

- A strong sentiment from the feedback was the need to distinguish ***production*** and ***use***. The ***production*** of statistics by Government departments and ONS requires rigorous collection and presentation of data, in line with the Code of Practice for Statistics. Once statistics have been published (ie ***produced***), they are available for ***use***, including by politicians. 'Production' can be thought of as an upstream activity, and 'use' as downstream. In the think-piece, we are focussing on the downstream element of 'use'.
- Although the paper focuses on downstream 'use', we should recognise that the way statistics are produced can raise risks of misinterpretation and hence be used in a misleading way. OSR frequently addresses issues with production, such as poor presentation or incomplete commentary, in our regular reviews of statistics. This work lies outside the scope of this paper.
- In thinking about 'use', we recognise that there is often a range of actors involved in presenting a claim about statistics: the Government body that

[1] Eliot Michaelson, Kings College London; Andreas Stokke, Uppsala University; Neri Marsilli, University of Barcelona; Alexandra Freeman, University of Cambridge; Jonathan Webber, University of Cardiff

produces and publishes the statistics; the communications team that presents information drawing on the statistics; media interpretation and summary of what is said; social media reuse of short segments of what is said; and many more actors. It is not OSR's role to intervene at all points in this chain. Our role is to focus on how prominent politicians take the statistics and use them in their own communications – for example, speeches, press releases, social media statements. There are other organisations, including Ofcom and press regulators, who consider the work of various media actors.

- In terms of 'use', there will always be a risk that too much weight is put on a particular set of statistics. As Paul Allin, chair of the Statistics User Forum, told us: "Statistics rely on precise definitions of things being measured, but result almost invariably in some imprecision on the measurement. Statistics are not strict accounts and may have confidence or error limits."

Intention not a helpful concept to guide our judgement

- Intention is not a helpful basis for guiding or supporting the OSR's judgements about misleadingness. Both regulators and philosophers agreed that deciding someone has intended to mislead is difficult, subjective, and likely to lead to unnecessary controversy.
- It is far better to consider likely impact on audience, rather than intentions of the speaker. This approach is consistent with that taken when judging misleadingness in other contexts, for example by the Advertising Standards Authority and Ofcom.
- Judging intention may be important to some people – for example, journalists wishing to understand and explain the factors behind particular decisions or arguments. But OSR is forming a view on the appropriate use and interpretation of the statistics, not judging the motivations and drivers of the person using the statistics.
- Although judging the intention of the speaker may not be the right approach for OSR, when considering materiality we will look at whether the use of statistics is significant – and one element in this consideration can be whether a particular use is repeated over time or part of a prepared communication (speech, political ad, etc). These factors will inform how we take forward a case – for example, how strongly we express any concerns we may have.

Aspects not covered by the initial think-piece

- Some specific issues arose that need further consideration:
    - There were some risk factors that the original think-piece did not consider, for example the use of incomplete statistical evidence (eg placing too much weight on early results of a new policy) or recency (eg placing too much weight on the latest data, even if changes the new data appear to show are not meaningfully different from past data).

- Many of the cases that OSR will deal with are relatively simple – for example, false statements that should be corrected, or use of unpublished data. In these cases, misleadingness would not be considered. In more complex or ambiguous cases where it is harder to reach a judgement, OSR would consider whether a statement has been misleading.

Use of statistics in political communication

- The think-piece is relatively silent on the role of intermediaries. In one of our conversations we discussed a scenario in which a speech that is carefully constructed, well researched and uses statistics appropriately, is summarised in a single soundbite in media reporting. It was suggested that the speaker may actually intend this outcome, knowing that a careful speech will inevitably be packaged into a soundbite that could be misleading.
- As noted above, in the case of media intermediaries, the OSR approach would typically focus on the content of the original communication, not the media reporting of it. In the same way that we can't assume the intentions of a speaker, it is similarly difficult to comment directly on the interpretation made by intermediaries. We can however give our view on the correct interpretation of the underlying statistics.
- One particular feature of political rhetoric, highlighted to us by Thomas King of the RSS Data Ethics section, is that different actors can draw widely different conclusions from the same underlying evidence. The point of democratic discussion is that different arguments are put forward; different narratives are presented; and different visions of good policy and the public interest are articulated. OSR's role is not to judge these different perspectives, nor to limit the use of statistics to support them. Instead, our role is more humble: we simply try to ensure that the statistics are used in a way that does not give a misleading impression of the statistical picture.
- It is not OSR's ambition to be an arbiter of political debate, nor would it be appropriate. Our role is to protect the role of statistics in public debate – that is, to ensure that their content and any caveats are respected in the way that they are used.

Evolving the think-piece

Based on the findings from above, we have evolved how we consider these questions, by downplaying intention, recognising complexity, adding in further risk factors, and being clearer on the circumstances in which it is relevant to consider misleadingness.

For simple cases which are about false statements these considerations are not relevant. (An example is provided in the annex in which a clear misstatement about education funding was brought to our attention and was

subsequently corrected). However, for complex cases, which are about the interpretation and weight put on statistics, these considerations are relevant. In all complex cases, we would use the core definition below to guide our judgment.

Old definition:

"We are concerned when, on a question of significant public interest, the way statistics are used is likely to leave a reasonable person believing something which the full statistical evidence would not support."

Updated definition:

"We are concerned when, on a question of significant public interest, the way statistics are used is likely to leave **audiences believing something which the relevant statistical evidence would not support."**

Approach:

Each piece of casework will be subjected to the same initial consideration, asking the following question:

> Is this a question of 'use' of statistics, or of 'production'?

If it is the latter, OSR will consider the issue in line with our interventions policy and the Code of Practice for Statistics, and look to address the question**:**

> We are concerned when, on a question of significant public interest, the way statistics are used is likely to leave **audiences believing something which the relevant statistical evidence would not support.**

In addressing this question, there are three aspects to consider:

1: The nature of the issue

It is important to start with the issue and the context. This will enable consideration of whether relevant audiences are likely to be misled about a particular set of statistics, and whether there is any evidence that they have been misled. [This is based on approach 2 from the original think-piece]

There could be a range of audiences, of course, ranging from technically knowledgeable specialists to the general public, and OSR should consider which of

these audiences is most relevant in considering the way the statistics have been used.

2: Risk factors

There are some recurring features of the way statistics are used that constitute risk factors – factors that can give audiences a different impression from that provided by the full, underlying evidence. [This is an extended version of approach 3 from the original think-piece]

The risk factors are:

1. There is selectivity of data points to support a claim which other data points do not support. (for example, from a time series)

2. There is selectivity of a metric to support a claim which other related metrics do not support (for example absolute figures rather than percentage or cash terms rather than real terms)

3. The language used does not fully represent the available statistics (for example implying the statistics represent a much broader or narrower definition than appropriate.)

4. There are methodological choices which lead to potential bias in the presented figures

5. No source or methodology is given, making it likely that a hearer could draw inaccurate conclusions about what the available statistics represent

6. Poor quality data is used, making it likely that the hearer will believe something which is untrue

7. There is an inappropriate choice of graph axis or data

8. The causality of a statistic is overstated, making it likely that the hearer will believe there is stronger evidence to support a causal link than exists

9. There is an error in the statistic used – for example, the figures for the wrong year are used to describe a change over time

10. ***NEW*** There is undue weight put on recent or new data

11. ***NEW*** There is too much emphasis on data that are incomplete. (For example, early results from a trial)

3: Materiality

Not all uses are as prominent as each other. It is important to consider the context of the use of statistics and ask the following questions:

- Is it a one-off or repeated use?
- Is it on a subject that the speaker has formal responsibility for?
- Is it part of a prepared speech or not?
- What is the public profile of the person using the statistics?

The answer to these questions will determine how significant the issue appears to be, with a one-off remark being less significant than a repeated use. [This is based on approach 1 from the original think-piece, but with no consideration of intention]

Next steps

Although we already employ many of the concepts in this think-piece in our ongoing work, it is not yet finalised and we will continue to explore how it operates in practice. We would also welcome further comment to guide future updates and improvements.

If you'd like to get in touch with us about this document, please email us.

Annex: Case examples

1. School funding 2020

The issue

A statement by the Conservative Party claimed that school funding had been increased by a minimum of £5,150 per pupil. A Labour Party MP, Wes Streeting, raised this with the UK Statistics Authority and said that the claim was false and misleading.

OSR's review

We reviewed the available data on school funding and concluded quickly that this claim was a mistake. The level of funding had increased to £5,150, but the increase itself was only £150. This analysis aligned with that undertaken by others, including Full Fact. The Conservative Party also removed the claim from its published material and deleted the Tweet.

Consideration of misleadingness

This was not a case of misleading presentation or use of statistics. It was more straightforwardly wrong. It was not therefore necessary to consider misleadingness to guide our judgement.

2. Crime statistics 2019

The issue

In a speech at the launch of the Labour Party manifesto launch, the leader of the Labour Party, Jeremy Corbyn, stated that "violent crime had doubled under the Conservatives' austerity programme."

OSR's review

There are two statistical sources on crime levels: crimes recorded by the police and crimes reported by the public in a survey. We have had long-standing concerns about the police recorded crime statistics and in 2014 we removed the National Statistics designation from the police recorded crime statistics because of quality concerns. Subsequent audits by HM Inspectorate of Constabulary and Fire Services demonstrated that quality of crime recording remained unreliable. Our long-standing public position has been that, for trends in most crimes, the ONS's crime survey provides a more reliable source than the police recorded crime statistics. While the police recorded crime statistics showed a trend broadly consistent with the Labour

leader's statement, the ONS survey showed no significant change in levels of violent crime.

Consideration of misleadingness

- The nature of the issue
  Although the statement could be defended on empirical grounds – there was certainly a time series that showed the trend being claimed – we considered that there was a risk that audiences could be misled by the statement, given its prominence in a prepared speech used to support a high profile political campaign. There was some media reporting of the claim too.

- Risk factors
  The following risk factors are relevant:

  - **There is selectivity of a metric to support a claim which other related metrics do not support:** Although this case is not really about cash or percentage presentations of data, the core point is that a claim was made based on the selection of a metric that other related metrics would not support: in other words, using the more reliable ONS survey would not support the same conclusion.

  - ***The language used does not fully represent the available statistics, for example implying the statistics represent a much broader or narrower definition than appropriate:*** The claim did not make clear that it was drawing on a specific definition (crimes recorded by the police).

  - ***Poor quality data is used, making it likely that the hearer will believe something which is untrue:*** The poorer quality of the police recorded crime data is well established.

- Materiality
  We concluded that this use of statistics was material – it was designed for repeated use in political campaigning and was part of a prepared speech by a leading politician. We therefore advised the UK Statistics Authority Chair to write to the leader of the Labour Party, and the letter was published on the UK Statistics Authority website.

3. Covid prevalence 2020

The issue

In a speech, the First Minister of Scotland, Nicola Sturgeon, claimed on July 3 2020 that the prevalence of Covid-19 infections was 5 times lower in Scotland than in England. The claim was repeated by other Scottish Ministers and in related news media.

OSR's review

We could not easily find any publicly available source that would support this claim. When we approached the Scottish Government, they provided an explanation based on comparing estimates from two models. The model for Scotland was sourced from Scotland's COVID-19: modelling the epidemic (issue no.6) 25 June and the England prevalence figure was sourced from modelling work done by the London School of Hygiene and Tropical Medicine, using a UK estimate as a proxy for England, but which was not readily accessible.

The Scottish Government then compared the upper prevalence rates published in Scotland's COVID-19: modelling the epidemic (issue no.6) 25 June and the Office for National Statistics' COVID-19 Infection Survey pilot: 25 June. This was done to corroborate the figures from the London School of Hygiene and Tropical Medicine.

Comparison of prevalence rates is not straightforward. If it is to be undertaken, the results and the uncertainties should be communicated transparently. We do not think that the sources above allow for a quantified and uncaveated comparison of the kind that was made. In future if such comparisons are made, we would expect to see sources made publicly available and a clear explanation of the limitations and associated uncertainty.

Consideration of misleadingness

- The nature of the issue
  We considered this to be a material statement with the potential to influence both people's understanding and also, potentially, their behaviour (eg willingness to travel).

- Risk factors

- ***There is selectivity of data points, for example from a time series, to support a claim which other data points do not support:*** We felt there was some element of this risk factor, in that the use of upper bound estimates were chosen, rather than a range of central estimate. This appeared to us to be a selective approach.

- ***The language used does not fully represent the available statistics, for example implying the statistics represent a much broader or narrower definition than appropriate:*** We concluded that the language was far too confident in describing different prevalence rates between the two countries,

and did not recognise sufficiently the inherent uncertainties in comparing COVID-19 prevalence rates. We would expect to see a clearer explanation of the limitations and associated uncertainty.

- ***No source or methodology is given, making it likely that a hearer could draw inaccurate conclusions about what the available statistics represent:*** This risk factor was strongly present. The absence of a clear source, and the lack of clarity on methodology, represented a significant risk factor.

- ***There is too much emphasis on data that are incomplete eg early results from a trial:*** This risk factor was also present, although perhaps less strongly than some of the other risk factors.

- Materiality
  We concluded that this use of statistics was material – it was designed as part of a prepared speech by a leading politician, based on advice and analysis from the Scottish Government. We also judged that this was potentially misleading, in conveying a stronger conclusion than the available evidence would support. We therefore wrote to the Chief Statistician of the Scottish Government to outline these concerns.