Office for Statistics Regulation

Misleadingness: A short thinkpiece

May 2020

Contents

Foreword – Jenny Saul	**3**
Introduction	**5**
Background	**8**
Approach 1: Materiality and Intention	**10**
Approach 2: Speaker and hearer: an audience-led approach	**14**
Approach 3: Case-led	**16**
Conclusion	**18**
Index	**20**

This is not a guidance document, but rather an exploration of the different ways that we as regulators could tackle misleadingness. We are eager to share our thinking and ignite debate on this topic and we invite you to share your views with us.

Misleadingness: not just ‘blue sky research’

Foreword – Jenny Saul

My interest in misleadingness arose out of my study of communication. For many years, I have been especially interested in the ways that we manage to communicate messages that go beyond what is strictly said.

In particular, I was interested in the fact that most (though not all) people seem to care a great deal more about *saying* something false than merely communicating it. This distinction is often captured by contrasting lying with merely misleading. We see the importance attached to it in the way that people will go to considerable lengths to mislead without saying something false. A classic recent example that I’ve discussed in my research is Bill Clinton’s famously misleading (but not lying) utterance, “There is no sexual relationship”.

In 2012 I published a book exploring what we can learn by thinking about the distinction between lying and merely misleading. One strand of the book examined what we can learn about language, and in particular about what is said. The other strand of the book explored why it is that we care so much about the distinction, and whether we should. In the end, I surprised myself by becoming convinced that it is just as bad to deliberately mislead as it is to lie.

Over the years since then, I’ve become increasingly concerned about many aspects of our public discourse, one of these being the ways that falsehoods are propagated. I was, as a result, very excited to be contacted by Ed Humpherson about developing a framework for assessing misleading uses of statistics. The process of developing this framework has been immensely enjoyable, and I think the results shed real light on varieties of misleading statistics and appropriate ways to think about and react to them.

The understanding of misleading that we eventually arrived at focuses on whether a reasonable audience would be likely to be led into a belief that’s unsupported by the available statistical evidence.

In assessing culpability, we look at factors like whether the speaker’s position means that they should have known better than to present information in this misleading way; and how important the misleading statement in question is. These are all factors that play an important role in the philosophical literature on the nature of deception and blameworthiness for it.

The clear framework developed will not, it’s sad to say, eliminate misleading use of statistics from public life. However, it will hopefully help a great deal in reducing it.

This framework will aid those assessing the misleadingness of ways that statistics are used. But, perhaps even more importantly, it will aid in communicating about

misleadingness, by clarifying both the nature and importance of various ways that uses of statistics can mislead.

My work on misleading was originally developed as ‘blue skies’ research, without thoughts of immediate application. But, thanks to an immensely fruitful collaboration, this research has helped the UK Statistics Authority to develop a framework that will help it to understand and communicate effectively about misleading uses of statistics.

To my mind, this is a wonderful example of the way that academic work can have a genuine and significant impact on matters of importance outside academia. The widely acknowledged problem of “fake news” is one of the most important threats to democracy.

The work of the Statistics Authority to combat misleading use of statistics is vital to combatting this, and I’m enormously proud to have played a small role in developing this framework.

Jenny Saul

Introduction – Ed Humpherson

In introducing this thinkpiece on frameworks for misleadingness, I want to make two points:

1. It's a very important part of our work at the Office for Statistics Regulation
2. The paper is deliberately exploratory. It is a thinkpiece – it's not designed to give a definitive view of what misleadingness means. It's designed to expose our thought processes and invite comments. It does so by exploring three alternative models for thinking about misleadingness.

I would also like to thank Jenny Saul, who has guided our development of this thinking and has been a tremendous source of wisdom and insight. If there is one meta-conclusion I draw from this piece of work it is this: work with philosophers more often!

1. Judging misleadingness forms an important part of our work at OSR

OSR is the regulatory arm of the UK Statistics Authority, a body established by the Statistics and Registration Service Act (2007). We are independent from government Ministers. We are separate from producers of statistics, including the Office for National Statistics. In line with the Statistics and Registration Service Act (2007), our principal roles are to:

- Set the statutory Code of Practice for Statistics
- Assess compliance with the Code of Practice
- Award the National Statistics designation to official statistics that comply fully with the Code of Practice
- Report any concerns on the quality, good practice and comprehensiveness of official statistics

The greater part of our work concerns how statistics are collected, produced and published, under the Code of Practice for Statistics. This focuses on the producers of statistics – essentially Government bodies – and in particular on trustworthiness, quality and value. We think of this focus as being upstream – looking at how information is collated and provided in the first place.

But we don't stop there at this upstream point.

For statistics to serve the public good it is not enough that they are collated, produced and published appropriately. They must also be used appropriately. So the Authority, supported by OSR, has always played a role in commenting on the use of statistics in

public debate. This can involve highlighting the correct interpretation of statistics that are controversial, and writing publicly to Government Ministers and other politicians about their use of statistics. We summarise this work annually in a review.

We are guided by a published interventions policy. This interventions policy sets out the criteria and approach that the Authority will take to making public statements on the use of statistics. The policy states that:

We are guided by a series of long-standing practices. The Authority will intervene if:

Official statistics are leaked before publication

The advice of professional statisticians is ignored

Official statistics are presented in such a way that they are liable to mislead the public or undermine the integrity of official statistics.

This thinkpiece explores the third of these aspects **– whether statistics are presented in such a way that they are liable to mislead.**

2. This paper is deliberately exploratory

Our approach in this paper is unapologetically exploratory.

It is a thinkpiece – it's not designed to give a definitive view of what misleadingness means. It's designed to expose our thought processes and invite comments.

This is a complex area, where careful judgement needs to be applied. We have extensive experience of making these judgements, including at relatively high speed during election campaigns. What this think piece aims to do is to guide people through our thought process – both to understand better what we do, but also to consider for themselves how to identify, address and mitigate the risks of misleading uses of statistics.

We want to explore ideas around misleadingness as an invitation to discussion, rather than setting out a definitive policy statement.

In line with our desire for the paper to set out our thinking fully, the paper is structured around three alternatives, each of which has limitations. It outlines each approach and then describes any flaws. We have structured it as a series of potentially flawed alternatives because we want to reflect our own thought processes in preparing the paper, and as an alternative to simply setting out a preferred option. In doing it this way, the paper follows the structure we used in our conversations with Jenny Saul.

The end point

This summary won't set out all the conclusions we reach – in line with the exploratory approach, we would prefer readers to follow the arguments, just as we did in discussions with Jenny.

But to give the end point first, we conclude that, when we judge misleadingness, what we are doing is putting the following consideration at the heart of our thinking:

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**.

Background

An important part of OSR's purpose is to protect the role of statistics in public debate. This is because, **for statistics to serve the public good, it is not enough that they are collected and published in appropriate ways. They must also be used appropriately.**

Failure to use statistics appropriately could lead to people being misled about aspects of society and policy; and could also lead people to lose confidence in official statistics. Both outcomes – people being misled and a loss of confidence – would mean that statistics were not serving the public good.

As a result, OSR will make interventions (or advise and support the UK Statistics Authority chair to do so). In these interventions, OSR typically highlights a misuse of statistics, and frequently articulates that a particular use is misleading or potentially misleading. The intervention takes the form of exercising public voice – making a statement or publishing a letter to the user of the statistic to highlight our concern and frequently to request a change in language or presentation. We do all this in line with principles and aims of the statutory Code of Practice for Statistics. [1]

So making judgements about misleadingness sits at the heart of OSR's role in protecting the role of statistics in public debate. There are lots of other organisations that operate in this space, including the news media and fact checking organisations. We have found that one organisation in particular provides a useful point of reference – the Advertising Standards Authority (ASA). The Advertising Standards Authority has extensive experience of judging misleadingness in the context of advertising, which it summarises in its Code – the CAP (Committees on Advertisting Practice) Code.

We also looked for philosophers who have thought about these issues, and very much liked the work of Jenny Saul on misleadingness (*Lying, Misleading and What is Said*, Oxford University Press, 2012). As a result, we met Jenny on several occasions in the course of 2019 and this paper reflects those discussions.

Three approaches

Here we explore methods that OSR uses to form judgements on whether particular uses of statistics are likely to be misleading; and on whether our intervention is justified.

It considers three approaches to misleadingness:

1. Materiality and intention
2. Speaker and hearer: an audience-led approach
3. Case-led

In each case, it describes:

- How the approach would work
- Issues, questions and flaws with the approach

Approach 1: Materiality and intention

One approach to making judgements on misleadingness would be to employ a matrix. This sort of matrix would weigh up importance (materiality) and aims (intention). It is adapted from risk management approaches (which assess risks by reference to impact and likelihood); and from our own Quality Assurance of Administrative Data (QAAD) which uses a risk matrix to judge the degree of assurance required over administrative data.

Figure 1 – Materiality/Intention

Materiality		
	High materiality, low intention: PRIVATE INTERVENTION TO REQUEST CLARIFICATION	High materiality, high intention: SIGNIFICANT INTERVENTION
	Low materiality, low intention: NO INTERVENTION	Low materiality, high intention: PUBLIC LETTER, BUT WITH LOW PROFILE

Intention

The two axes on this matrix are:

Materiality – how important is the misrepresentation of statistics? We would distinguish between:

- a statement on a matter of high public interest and debate – for example, a topic of a major Parliamentary debate or one that is prominent in an election campaign – and one that is of lesser interest
- a statement that is slightly wrong from one which is significantly wrong, including the extent to which the underlying statistics are themselves clear and straightforward to interpret
- a statement that is made by a prominent figure in public life from one made by someone with a less formal role

and

Intention – although sometimes hard to determine, we would consider factors like:

- whether the statement was a core part of an argument and was published after a review process and whether it was repeated; and whether the person is someone who should be expected to know the subject area well (e.g. a Minister or a shadow spokesperson on the topic).
- whether the statement is made in passing or is part of an ongoing communication campaign or policy debate. For example for a scheduled debate in Parliament or Committee appearance, participants are likely to be fully briefed by their respective teams, including on the available statistical evidence. So a misstatement here is much more likely to be a result of active choice than a passing comment. The key issue here is whether it is legitimate to expect someone to have done all their research or not.

The matrix would help us to determine whether to make an intervention and if so what kind. Compare two examples:

Example a: A politician responsible for the justice system appears on a TV news programme and is asked about a broader area of policy outside her formal responsibility (e.g. wider economic policy).

She makes an incorrect statement about what the statistics show (e.g. level of economic growth is lower than she states). She does not repeat the statement and no-one else re-uses it. **We would conclude LOW MATERIALITY and LOW INTENTION – and make no intervention.**

Example b: The same politician responsible for justice issues a statement, and then a speech, proposing a radical overhaul of the courts system.

The statement and speech contain statements about the historic performance of the court system which are selective and omit crucial contextual information about the volume of cases passing through the system.

We would conclude **HIGH INTENTION** – because it was clearly a prepared and thought-through line – and would judge Materiality by the extent to which the statement was incorrect and the relevance of the statement to the overall argument. **Either way, this clearly would require a public intervention.**

This matrix provides a simple and intuitive way of gauging the strength of an intervention.

But it is flawed in **three** respects:

- Intention is relatively subjective and largely determined by circumstantial factors e.g. how often is something said; how "prepared" is that statement. It's worth noting that in its CAP Code [2], the ASA explicitly rules out an intentions focus: the ASA "will rule on the basis of the likely effect on consumers, not the marketer's intentions."[2]
- The matrix assumes that it is easy to determine what is misleading in the first place (e.g. that we know that something is a problem and what we need to do is to determine intention and materiality).
- More fundamentally, the concept of materiality is being made to do a lot of work, covering a range of aspects around the statement, the context, the person who makes the statement and so on. It could well prove difficult in an actual case to reduce these various categories to a single categorisation of materiality.

The solution to this third flaw is to decompose the single criterion of materiality into a series of independent questions of impact:

- The role of the person making the statement (should they be expected to know the subject)
- Degree of incorrectness
- Context – the question being asked and/or the argument being made
- Degree of harm – how could an incorrect statement of or interpretation of the statistics lead to people to make harmful decisions (to themselves or others)

We would then form a judgement against these questions independently and not reduce them to a single concept of materiality, and form our judgement about how to intervene by their combined effect, alongside the intention questions. But this is, of course, all rather complicated.

Approach 2: Speaker and hearer: an audience-led approach

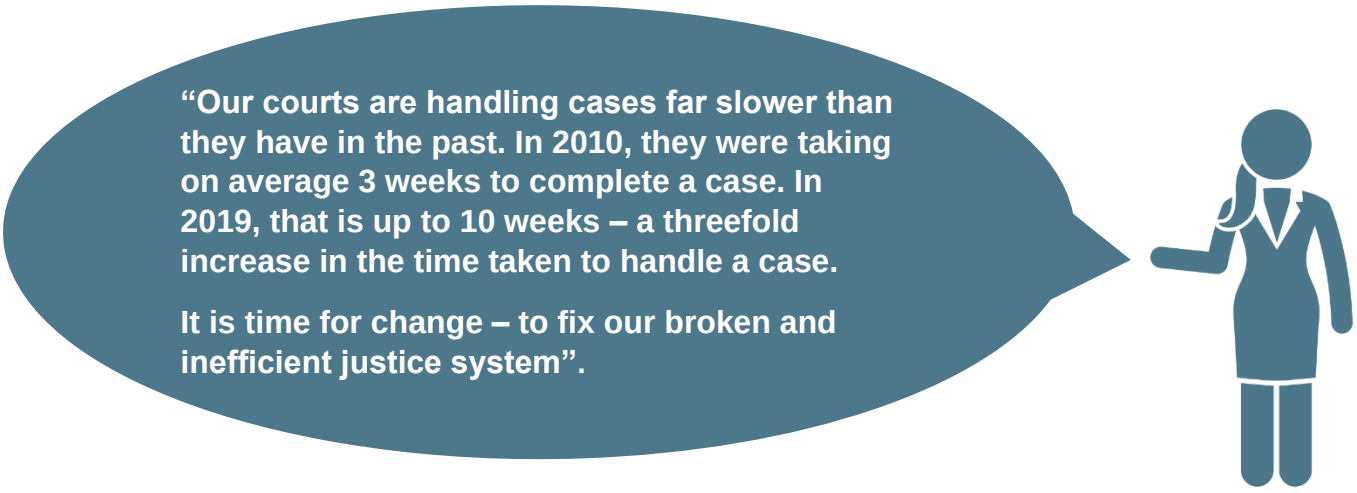
An alternative approach is to focus less on materiality and intention, and focus more on the roles of speaker, hearer and the interaction between the two.

In this context, the speaker is the person or organisation using statistics. It could be a politician; a Government department; political party; or other actor (e.g. a charity or a business). The speaker disseminates a statement that includes statistics.[3]

The hearer is the audience. It depends on context, but typically we would think of the audience as being members of the public who hear or read the statement. It could however be a more specialist audience in some cases. Audiences can be wide and varied – and in a social media context, statements can sometimes travel far beyond their originally intended audience.

For something to be misleading, **the hearer must actually draw a false conclusion** – that is, a) they conclude something that is not supported by a correct reading of the available statistics; and b) if they were to receive correct information, a reasonable hearer would form a different view. This approach is recognised in the ASA's thinking [4] and the focus on the conclusion drawn by hearers is one of the strongest influences of Jenny Saul on our thinking. And both conditions need to be satisfied – that someone concludes something that is not supported by the available statistics; and that a correct reading would lead to a different conclusion.

To take an example, imagine that in the Justice example above the speaker is the Minister for Justice, who says something like this:



The hearer could conclude that that the threefold increase indicates gross inefficiency and that it is indeed time to change.

But imagine the statistics were in fact stated in a misleading way – because the increase from 3 weeks to 10 weeks in fact is a result of a measurement change. Prior to 2017 (let's say) only time in court, not elapsed time, was counted – so a delay or adjournment stopped the clock. After 2017, the statistics measure elapsed time regardless of adjournment. Putting them on a like for like basis, the time for a court case was 9 weeks in 2010 and 10 weeks in 2019.

Once they had learned about this change, a reasonable hearer might well then change their view – that the case for change is not strong at all. And that they were misled by the comparison of 3 weeks and 10 weeks.

So there are in effect two tests going on here:

1. Is the statement a reasonable reflection of what the statistics actually say? (answer – no, because the statement does not acknowledge the method change)

2. Is a hearer misled i.e if they were to receive the correct information would they form a different view? (answer – probably yes, because the correct interpretation is so different from the initial impression).

OSR could sensibly split our judgement in two along the lines of these two tests, asking first whether the statement is a fair reflection of the statistics; and if not, forming a view as to whether the misstatement is sufficient to shift a reasonable hearer's view.

But there is a problem here too. Rather as in the first framework, where the problem is the difficulty of knowing the subjective intentions of the speaker, here the problem is knowing the subjective experience of the hearer. Short of doing large scale surveys of the public on every statement, including before and after tests, it would be very difficult to conclude what a hearer understands or infers. [5]

Approach 3: Case-led

This then leads into a third framework which is based on past OSR work.

To recap:

- We can't reliably know the intention of the speaker (although to be fair we can sometimes infer it) and have to form complex judgements on importance that are difficult to boil down to a single concept like materiality
- We can't always reliably know the impact on the hearer (although survey evidence can provide some evidence to help us infer it in some circumstances)

But what we can do is highlight actions on the part of the speaker that, in our view, raise the likelihood of misleading hearers (whether or not they actually do is of course an empirical question that would be difficult to research in every case).

The history of the cases that we have worked on furnish quite a lot of examples of actions on the part of people making statements that have caused us concern. They are set out in Figure 2 below.

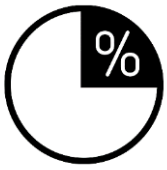
The common theme is that these issues relate to the ways in which speakers or organisations present their statements. They are almost all self-evidently problematic too – they do not require subjective judgements about speaker intention or hearer impact.

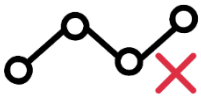
They range from data presentation – truncated axes – through to misdescription and implying causality where none is evident in the statistics.

And they all have a further feature in common – they all relate to ways in which the speaker or organisation making the statement misrepresents what the available statistics say (so that a hearer might think that the statistics say something that they in fact do not).

Figure 2: Issues that have caused concern in the past

- There is selectivity of data points, for example from a time series, to support a claim which other data points do not support
- There is selectivity of a metric, for example absolute figures rather than percentage or cash terms rather than real terms, to support a claim which other related metrics do not support [6]
- The language used does not fully represent the available statistics, for example implying the statistics represent a much broader or narrower definition than appropriate
- There are methodological choices which lead to potential bias in the presented figures
- No source or methodology is given, making it likely that a hearer could draw inaccurate conclusions about what the available statistics represent
- Poor quality data is used, making it likely that the hearer will believe something which is untrue
- There is an inappropriate choice of graph axis or data
- 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
- There is an error in the statistic used





The advantage of this approach is its pragmatism – it is based on observable actions by people and organisations making statements and is heavily grounded in existing OSR cases. This pragmatism also provides space for the kind of hearer/audience impact that is central to some notions of misleadingness (and particularly important in Jenny Saul's work).

Conclusion

This short note has set out three potential options for thinking about misleadingness. It has examined each and found limitations to them. So it is probably best to proceed by blending the three approaches:

- **Approach 1: Materiality and intention** - Using the intention/impact approach to form a judgement on the scale of OSR intervention
- **Approach 2: Speaker and hearer: an audience-led approach -** Considering whether there is evidence that people actually have been misled (while recognising that this sort of evidence is unlikely to be available very often)
- **Approach 3: Case led; factors that in our view are likely to mislead**- Adopting a simple set of criteria based on the producer acts described in Figure 2 above – like inappropriate choice of graph axis or data.

And there is an overarching principle that emerges from this discussion. At root we are not checking facts or policing debate. We are protecting a fair understanding of statistics.

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 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**."

Blending the three approaches

Approach 1: Materiality and intention

Use the materiality and intention matrix to form a judgement on the scale of OSR intervention (Figure 1)

Approach 2: Speaker and hearer: an audience-led approach

Consider whether there is evidence that people actually have been misled (while recognising that this sort of evidence is unlikely to be available very often)

Approach 3: Case led; factors that are likely to mislead

Look at specific factors that in our view are likely to mislead. Adopt a simple set of criteria based on the producer acts described in Figure 2:

- selectivity of data points or metric
- using poor quality data
- inappropriate choice of graph axis
- using language to imply the statistics represent a broader or narrower definition than appropriate
- choices which could lead to bias
- overstating causality
- not providing the source or methodology behind the data
- using statistics that are wrong or inaccurate

Index

[1] The UK Statistics Authority sets and upholds the statutory Code of Practice for Statistics. It can be found on the Authority's website at https://www.statisticsauthority.gov.uk/code-of-practice/. The Office for Statistics Regulation (OSR) is the Authority's executive office for regulation.

[2] CAP Code 3- misleading advertising, https://www.asa.org.uk/type/non_broadcast/code_section/03.html

[3] The terms "speaker" and "hearer" are used broadly here, to encompass not just spoken words that are broadcast and heard; but written words that are published and read; or posted and viewed; or signed language.

[4] CAP Code 3: "The ASA will take into account the impression created by marketing communications as well as specific claims. It will rule on the basis of the likely effect on consumers…" https://www.asa.org.uk/type/non_broadcast/code_section/03.html

[5] However, this risk is somewhat migitated, at least in the case of the ASA, because under the Unfair Commercial Practices Directive (the EU's and UK's underpinning 'misleadingness' law), the 'average consumer' test is not a statistical test and decision-makers (including us) have to exercise "their own faculty of judgement, having regard to the case-law of the Court of Justice, to determine the typical reaction of the average consumer in a given case". This is the effect of the case law (see Gut Springenheide and Estée Lauder cases) and is reflected in Recital 18 of the Directive itself.

[6] The ASA refers to omission: "Marketing communications must not mislead the consumer by omitting material information. They must not mislead by hiding material information or presenting it in an unclear, unintelligible, ambiguous or untimely manner." CAP, 3.3. https://www.asa.org.uk/type/non_broadcast/code_section/03.html