What does OSR think about misleadingness?

OSR has a new working definition of misleadingness:

We are concerned when, on a question of significant public interest, statistics are used to communicate a descriptive statement that the wider relevant statistical evidence would not support, despite otherwise being an accurate statement.

The rest of this article explains how we reached this position and what it means for how we make casework judgements.

As we continue to develop our thinking on misleadingness, we remain open to feedback and comments. If you would like to get in touch with us about misleadingness, or anything else for that matter, please contact regulation@statistics.gov.uk.

Misleading statements

What does it look like for statistics to be used in a misleading way? The case study below describes one such instance in which OSR intervened.

"Debt is falling"
The case: Towards the end of 2023, we were contacted by an MP who raised their concerns about claims that the then-Prime Minister Rishi Sunak had made on different occasions and on different platforms that "debt is falling" and that "we have indeed reduced debt". The MP was concerned that these claims were not supported by the Office for National Statistics' latest Public Sector Finances statistics and were therefore misleading. **Our considerations**: As with any case, we looked into the concerns that were raised with us to understand not only if the claims had been made as stated, but also to see if we shared the same concerns and would therefore want to intervene. **Our findings**: Our findings were not as clear cut as you may expect. We were clear that the Prime Minister had made the claims as reported to us, but the data were less obvious. An upward revision to Gross Domestic Product had reduced the debt ratio in the most recent months of the claim. However, debt remained at a record high of 97.8% in October 2023. In addition, the Office for Budget Responsibility's March 2023 economic and fiscal outlook showed that public sector net debt as a % of GDP was forecast to fall between 2024–25 and 2028–29. **Our intervention**: In our response to the MP, we said the veracity of claims of this sort – with the brevity typical of a social media post or a statement made during a parliamentary debate – often depend on the wider context, as well as how they would be stated in a fuller form. Changes in public sector debt also might depend on the following factors:

- How you define debt? Are you including or excluding public sector banks, or the Bank of England? Are you referring to debt in cash terms, or as a % of GDP?
- What time period are you referring to? Do you mean over the last month, year or 10 years? Or are you referring to forecasts for the current year or longer?
- Was the intention to refer to overall changes or changes from policy action?

We judged that the average person in the street would probably not have interpreted the Prime Minister's claims in the way that his office explained them to us. They would likely have assumed that he was claiming that debt was already falling or that the government's policy decisions had lowered it at the fiscal events – neither of which is the case. This was a clear source of confusion and may have undermined trust in the government's use of statistics and quantitative analysis in this area.

When speaking about public finances and making claims of this sort, intelligent transparency demands that ministers, other senior politicians, departments and political parties ask themselves how someone with little specialist knowledge is likely to interpret a particular claim. They should explain their claim as clearly as possible, particularly where technical detail is left out and misinterpretation could result.

To see more information on this case, please read the MP's letter to the UK Statistics Authority (UKSA) regarding the debt claim, and UKSA's response.

OSR's work on misleadingness: the story so far

Misleadingness, how to define it and what this means in our context as a statistics regulator is something that we routinely come back to as the Office for Statistics Regulation. It's an area that we often consider within our casework function when we make judgements on how statistics and data have been produced, used or communicated.

In many of our casework judgements, categorising something as misleading is more nuanced than looking at whether a statement is supported by a dataset. As illustrated in our case study. How we define misleadingness, and our thinking around misleading cases, are areas that we will continue to develop as our casework function and regulation role continue to evolve.

Our work on misleadingness started with a think piece in May 2020 and a follow-up think piece in May 2021 with input from Jenny Saul, a philosopher who has written and thought extensively about misleadingness.

Our first think piece explored three approaches to misleadingness: materiality and intention; an audience-led approach, considering the speaker and hearer; and a case-led approach. This culminated in the 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."

This work took us beyond our usual area of expertise, so we were keen to invite feedback on our thinking. A strong sentiment from the feedback we got was the need to distinguish between the production of data and statistics and their use. We also found that speaker intention (what a person hopes to achieve by making a statement) was not a helpful basis for guiding or supporting our judgements about misleadingness. Determining whether someone has intended to mislead is difficult, subjective and likely to lead to unnecessary controversy.

This led us to our second [think piece](#), which redefined misleadingness in our context as "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.**"

This definition helped us judge misleadingness more consistently when considering casework concerns. But we began to see a trend in cases raised with us where the data in question were error free but not supported by the wider evidence base. This felt like a gap in our thinking, which created a useful starting point to review our approach to defining misleadingness.

In 2023, Dr Jenny Saul introduced us to a student of hers, Kyle Adams, who was interested in working with us as part of his PhD. We used this as an opportunity to revisit misleadingness and see if our thinking had changed. You can read about Kyle's perspective of working with us in [this blog](#).

With a focus on OSR's context, Kyle's research posed the following questions about our definition:

- Might some parts of our definition be better understood as guidelines for when to intervene, rather than as a definition?
- Is it appropriate to base a definition on an audience's beliefs?
- What do we mean exactly by "audiences"?

To define or guide?

Kyle first questioned the inclusion of "significant public interest" in the definition. He suggested that if our definition was to be taken as a practical guideline for which cases merit an OSR intervention, then this inclusion was appropriate. If, however, our definition was meant to characterise what misleadingness is, then "significant public interest" didn't seem appropriate or relevant.

Moreover, if statistics were used in a way that was likely to leave a small audience with a wrong impression, that would still be misleading – despite not meeting the criteria of "significant public interest".

Overall, he concluded that our definition was more a guiding principle for when OSR may intervene. Having considered his analysis, we agree with this conclusion. As a result, this has been added to our casework [FAQs](#) to ensure transparency around our casework process.

Conflicting beliefs

Kyle questioned the use of the criterion "likely to leave audiences believing something which the relevant statistical evidence would not support" in our definition of misleadingness.

A hypothetical case where this criterion might cause problems is as follows. Suppose that statistical evidence showed that offices that had pink sticky notes had higher staff turnover than offices who did not have pink sticky notes. One belief resulting from this is the **descriptive** one: that pink sticky notes cause staff turnover.

Some people may believe that decreasing staff turnover is a "good" thing and will therefore conclude that offices should stop buying pink sticky notes. This is a **normative** conclusion – a belief that depends on a consensus view of what constitutes a "good thing". But a different audience might believe that higher staff turnover is a necessary feature of a well-functioning office. This audience would end up with the same descriptive belief (pink sticky notes cause staff turnover), but a different normative one: offices *should* purchase pink sticky notes. In short, audience beliefs include both descriptive and normative beliefs.

Normative and Descriptive
In our discussions with Kyle, he often focused on differences between regulatory responses based on normative beliefs and descriptive beliefs. A normative belief, is a belief about what *should* be, and relates to evaluative standards (or norms). A descriptive belief is a belief about what *is*, and is generally independent of these kinds of standards. For instance, "children should eat vegetables" is a normative claim, whereas "many vegetables are rich in nutrients" is descriptive.

In cases where different audiences have conflicting normative beliefs, it can be confusing to determine which is best supported by statistical evidence. The statistical evidence might not support either belief on its own, but it also doesn't undermine either belief.

So where does that leave us when our guiding principle for how we intervene asks us to consider what "audience beliefs" are unsupported?

Kyle concluded that the criterion of support for audience beliefs was inadequate, and suggested limiting the kinds of claims that OSR considers. The hypothetical Post-it note case outlined above, coupled with OSR's official neutrality about what constitutes the public good, suggest that it might be appropriate for OSR to focus only on descriptive beliefs. OSR should be concerned with the wider body of statistical evidence to decide if a claim being made is misleading or not. We support this conclusion, as our role as a regulator is to safeguard the responsible use or statistics in public debate -- not to judge what is in the public's best interest.

Audience and reasonable people

Kyle then turned his focus to what we mean by "audience" and whether what may be misleading to some, may not be to others. For example, when a physicist says that a vacuum sucks in the surrounding air, an audience of other physicists would understand this to mean that surrounding fluid pressure pushes air into the vacuum.

However, a non-academic audience is less likely to understand the different mechanics at play.

This led to a conversation about what constitutes an audience and if the focus for OSR should be something more representative of the general public, which we decided could be defined as "a reasonable person".

How the general public would interpret a claim is important to us; our vision, after all, is statistics that serve the public good. We find it helpful when forming casework judgements to apply the lens of what a reasonable person may understand about a statistical claim. It is an important part of our intervention process, impacting how we communicate our judgements.

Overall, we agreed that our definition of misleadingness should focus on the relationship between the statement being made and the broader body of statistical evidence. This approach means our definition is more rigorous. The focus on the view of a "reasonable person" allows us to articulate our judgements in an accessible way.

OSR's approach to misleading vs. incorrect use of statistics

The final question Kyle asked us was around how we intervene when something is misleading compared with something is categorically wrong. A claim that is misleading may leave an audience believing in something which the relevant statistics don't support, but so would a claim that was completely incorrect.

It is important to highlight here that our role with interventions is to support the role of data in serving the public good. Our priority will always be on explaining what the data, highlighting what they do and do not show, rather than focusing on the communicator who may have used statistics incorrectly or misleadingly. We are not the arbiters of public debate.

Kyle suggested that our current definition of misleadingness could be used for both a misleading claim and an incorrect claim.

Redefining misleadingness

What does all this mean for our definition of misleading? Well, as we have already highlighted, it means that we have a new definition:

We are concerned when, on a question of significant public interest, statistics are used to communicate a descriptive statement that the wider relevant statistical evidence would not support, despite otherwise being an accurate statement.

Each aspect of this new definition plays a part in our judgment of misleadingness.

In order to understand what action to take or if we intervene at all, we need to ensure that there is significant public interest. Whilst we would discourage misleadingness at any level, it may not be proportionate, or aligned with our vision for statistics to serve the public good, to intervene on cases that lack public interest. In fact, by intervening, we may increase the opportunity to spread misinformation than if we had not intervened at all.

Statistics on their own are likely to be harmless. It's only when they are used to communicate a descriptive statement that the wider relevant statistical evidence doesn't support that misleadingness creeps in. Wider relevant statistical evidence is important here as it could encompass anything relevant – from statistics to guidance or information.

For example, the Home Office publishes statistics on illegal migration. These statistics record, among other things, the number of asylum applications awaiting an initial decision. The total number of applications awaiting an initial decision on 31 December 2023 was 98,345. This statistic alone cannot be misleading. However, when the then-Prime Minister claimed that he had cleared the backlog of asylum decisions by the end of 2023, the wider relevant statistical evidence did not support his claim: an ordinary understanding of "cleared the backlog" would suggest no cases are awaiting an initial decision. However, the Prime Minister's claim was supported by a Home Office press release and ad hoc statistics highlighting important relevant wider statistical evidence showing that the Prime Minister's commitment of clearing the legacy asylum backlog had been delivered *if* you accepted the context that all cases had been through some sort of initial review.

The final element of our definition is often key in deciding if a case is misleading. For a case to be misleading, statistics have to have been used to communicate a descriptive statement that the wider available statistical evidence would not support. This is different from something being materially wrong. As an example, if interest rates fell from March to April but overall remained higher than at the start of the year, it would be correct to say "interest rates reduced from March to April". It could be misleading to say "interest rates are falling" based solely on one month's data, because the wider available evidence does not support this claim. And it would be materially wrong to say "interest rates are at an all time low".

Our revised definition for misleadingness tightens the focus on what is and is not misleading. The additional context of a reasonable person's belief will be helpful when explaining our casework judgements for misleading uses of data.

We are looking forward to putting this new definition to the test in our casework. We would like to thank Kyle for his questions and thinking on this topic, and Jenny for introducing us.