

Language and public policy

Proposed standards
for the use of language
in social policy documents,
and social science research;
a guide for funders of research;
researchers; subjects of research;
policy makers; and the general public.

Matt Berkley

Draft, 6 January 2006

33 Howard St
Oxford OX4 3AY
England
+44 (0)1865 726768
www.mattberkley.com
matt@mattberkley.com

Contents

Summary	3
Purpose of this document	4
Background.....	5
Preliminary note.....	6
Proposed standards for policy and social science	7
1. Consider reliability of data for purpose.	7
2. Think about the reliability of conclusions.	8
3. Distinguish data on samples from inferences on populations.	10
4. Distinguish reports from history.	10
5. Distinguish population trends from trends for people.	11
6. Distinguish spending from income.	13
7. Distinguish spending from consumption.	14
8. Distinguish consumption from adequacy.	14
9. Distinguish income from profit.	15
10. Distinguish prices from relevant prices.	16
11. Distinguish prices from cost of living.....	17
12. Distinguish consumption gains from material gains.	18
13. Distinguish incidence from prevalence.	19
14. Distinguish prevalence from degree.	19
15. Distinguish material conditions from judgements on well-being.....	19
16. Imagine real people.	20
17. Look at meaningful groups of people.....	20
18. Distinguish “statistical significance” from importance.....	21
General note on the standards	22
Suggested examination questions.....	24

Summary

This document presents an initial list of proposals for the use of language for

- a) some public policy goals and
- b) some reporting in social science.

At the same time as proposing standards for researchers, funders and policy makers, the document aims to provide consumers of research output with ways of questioning past claims about poverty and prosperity.

The document mentions potential, and some actual, errors in social science.

The impetus was the author's interest in current methods of economic analysis. In the process he developed an urge to understand the relationships between

- a) current theory in social science
- b) predominant practice in social science and government and
- c) what he thought he might like as government aims and progress indicators if he were among those called "extremely poor".

In the process, also, he came to believe that thinking about the extreme case sometimes reveals things useful to understanding principles.

Purpose of this document

The aim of this document is to make clearer some aspects of government policy, and some aspects of social science reporting. The main focus is on economics.

Many people die of starvation in a world where resources are adequate. At the same time, people differ as to what they mean when they say “poverty has got worse” or “poverty has got better”.

One part of a solution to the starvation problem may be for professional organisations, institutions and funders, including government agencies, to adopt common language.

If this happened, then discussions about what has happened in the past might be more constructive.

Some questions related to social science are fundamentally subjective. That is one reason why the document is aimed at a wider readership than academics.

One aim of this document is to clarify which are scientific matters, which are matters of opinion and which are moral matters.

The distinctions in this document may help clarify some claims concerning human poverty and prosperity.

One reason for adopting the proposed standards is this: Looking at what lies behind the language of social scientists may help solve some puzzles in international statistics.

The purpose of this document is to further the following aim: that public institutions adopt standards for clearer language.

In 2000 the present author raised a fundamental flaw with professors of economics: if the “poorest” die, the figures appear to show a better outcome for them. The history of that issue among academics since that time leads the author to think that it is not enough for academics to discuss issues. What is needed are rules forbidding social scientists from making statements for which they clearly do not have evidence and which may have important social impacts.

At present, the author knows of no political party which endorses such rules in relation to the use of economic data.

Background

In 2000 it struck me that the debate over global poverty needed better statistics. I came to realise it needed clarity about existing statistics.

I was a trustee of a family trust. It occurred to me that the trust might help provide statistics for the debate on global poverty. I also thought it might help bring academics together with campaigners.

I then read an economic policy document taken seriously by newspapers and the UK government, which contained elementary errors. This situation seemed to indicate widespread errors of reasoning in economics.

Around the same time, the heads of several campaigning and research organisations told me that it would be a good thing if there were a think tank on global poverty.

Perhaps I was aiming to clarify two things:

- 1) how existing policy and research methods related to what I might want for myself in the situation of the “poorest”;
- 2) the evidence for the most influential claims about policies and poverty.

Some of the issues go back a long way.

For example, the tradition in macroeconomics (large-scale economics) has been to assume that if the incomes of the poorest rose 1% they had a 1% benefit.

Adam Smith noted a difference between the inflation rate for working-class people and the national rate in 1776.

He also wrote about needs being a factor in prosperity: he did not advocate the idea that resources measured either prosperity or poverty without reference to need.

Another example relates to adding up the numbers. Economists often write about utility - meaning the consequences for people. The idea of calculating “utility” or gains in well-being to people goes back to the philosopher Jeremy Bentham. His form of “utilitarianism” was about “the greatest good to the greatest number”. He included duration of pleasure as a factor in his “calculus” of consequences. But economists for some reason almost all ignored this in making claims about poverty.

It appears to have been common in macroeconomics to claim to know economic benefits to the poorest people without thinking about prices, needs or survival rates.

There is disagreement over what prosperity is. There is also disagreement over what might be evidence for it, and what have been trends in prosperity or its inverse, poverty.

It may be that clearer language would help resolve some aspects of disagreement.

Preliminary note

The standards are not meant to override common sense.

What counts as common sense can be a subjective matter.

Science is about approximating to the truth.

The aim of this document is to further that enterprise.

Proposed standards for policy and social science

1. Consider reliability of data for purpose.

Suppose a social scientist says they have data for something (poverty, hunger, survival, and so on). How do you go about deciding whether to believe them?

One type of question to ask is about reliability.

Do the numbers come from samples of the population, or everyone?

Usually, numbers come from samples.

This kind of thing may seem at first complex to analyse, but if you apply imagination you can think of some relevant aspects. After all, the scientist has obtained the data from somewhere. Often, the numbers are not actually about the thing the scientist mentions, but about what people say.

So if a researcher says “people are 2% happier in country X than in Y” behind the claim may well be something like this:

“In each country, workers asked 10,000 people (out of a country of 100,000,000) how happy they were.”

Now, there will be many details. But the principles are often easy to understand:

- 1) you can't really measure happiness - you don't really know what is going on in people's heads or in any other part of their body, and sometimes they don't either;
- 2) for cost reasons, the sample may be one in 1000 people or fewer
- 3) some people won't tell the truth
- 4) people will feel different at different times
- 5) the questions might seem different in different languages
- 6) the workers might have left out people they couldn't find, or who wouldn't answer
- 7) some of the above factors might have skewed the results

and so on.

This example is not meant to reflect the reality of research on what people say about happiness. It is to illustrate the principle that you can think about what social scientists tell you, and the principle that it's not magic - it's just people finding out about things by methods that are possible in real life.

My personal belief is that in the case of what some people call “extreme poverty” a strange kind of psychological process has happened whereby the “non-poor” can have a kind of short-circuit of imagination. Perhaps this is common in other areas of work as well. What I have consistently observed is that some people paid highly to be experts in “poverty” often fail to

show evidence of having thought about some of the most basic aspects of real life. Some of these aspects are mentioned below. They include questions about the reliability of data.

Researchers should estimate reliability of data for the specific purpose, giving reasons.

Estimates should be in the context of the specific statistical tests.

Data may be reliable enough for a simple test but not for a complex test.

This may sound complex, but the principles are simple: think about

- a) what you are comparing and
- b) how your confidence in the data might stand these comparisons.

Example:

“We estimate that
in the context of
comparing outcomes under policy X with outcomes under policy Y,
taking into account
the number of countries in each category or being correlated,
the sample sizes,
the survey coverage,
the possible gaps in data,
the number of time periods,
[...and any other relevant factors...]
the likelihood of our figures being right to within a% is b.”

Note that the unreliability of a series of inferences is multiplied (see below).

2. Think about the reliability of conclusions.

Researchers must consider the reliability of inferences, giving reasons.

The unreliability of a series of inferences is multiplied.

For example:

If

there are two steps in my argument

and

each has 70% probability

then

I am probably wrong.

$$70/100 \times 70/100 = 4900/10000 = 49\%.$$

Suppose survey data are on answers about spending. This is mostly the case in real life.

If an economist is asked by a politician to say something about how well or badly poor people did, here would be some necessary inferences.

From

i) "what people sampled in 1990 and 2000 said they spent"

to

ii) "what representative samples of people in 1990 and 2000 said they spent"

to

iii). "what poor people spent in 1990 and 2000"

to

iv) "trends in spending for real poor people over time
(taking demographic change into account)"

to

v) "trends in income for poor people"

or "consumption trends for poor people"

to

vi) "consumption adequacy for real poor people"

to

vii) "material gains for real poor people"

to

viii) "gains in well-being for real poor people".

Necessary assumptions would concern at least, not necessarily in this order:

- a) sampling adequacy
- b) truthfulness and memory of respondents

- c) demographic change (which determines food needs)
- d) workload (which determines food needs)
- e) survival rates
- f) a theory of how well-being relates to spending
- g) the relative value to people of assets and consumption

and so on.

In order to understand the inferences it is perhaps important to be clear about different kinds of statements.

3. Distinguish data on samples from inferences on populations.

The statement

“People in our sample in country P had rises in X”

is not the same statement as

“In country P, X rose”.

Researchers should note non-responses and difficulties in obtaining random samples. The aim here is to exclude the risk of error through sample bias.

Note: The “rich” and the “destitute” may not be reachable in surveys. If the destitute are unreachable, it is not clear how an economist can have data on the poorest.

Note on units chosen: If the units chosen for comparison are countries, reasons should be given as to why these countries might constitute representative samples of all relevant countries.

Similar considerations apply to time periods.

4. Distinguish reports from history.

What people said does not tell you what they did.

Economic data on most people are on people's answers to questions about spending.

In many areas of life, answers people give may not be true.

These areas of life include what they spent, earned, ate, drank, used, or acquired.

Reasons relate to

- a) honesty
- b) self-deception
- c) memory
- d) mathematical ability.

It is a mistake to describe answers as measured quantities without good reason.

5. Distinguish population trends from trends for people.

Reason:

To minimise risk of damage to people from confusion of:

- a) trends for people
- and
- b) changes in aggregates for populations

Example:

"the average rise [for people] was x%"
versus
"the [population] average rose by y%".

Demography axiom

It is not possible to aggregate outcomes from statistics solely on the living.

Note

Statistics on survivors are selective. Aggregate outcome statistics include those who die.

Statistics on survivors do not yield data on what happened to people during a period.

Nor do they yield data on what happened to the survivors over the period.

The statement
“The average was x% higher in 2000 than in 1990”
does not tell a researcher either that
“people had rises of x%
or that
“survivors had rises of x%”.

Examples of what are strictly speaking errors:

a) Any inferences as to aggregate trends for people from United Nations Millennium Goal indicators for hunger, poverty, education, AIDS, water.

These indicators were in terms of population proportions. They are therefore not indicators of aggregate progress for people.

Numbers of living people depend on births, deaths and migration as well as trends for individuals.

Proportions of living people either side of a line described in terms of proportions of people depend on births, deaths and migration of people on each side of the line as well as on individual trends.

b) Policy assessments from macroeconomic statistics.

Treatment of population statistics as statistics on people has been the dominant tradition in macroeconomics.

Exceptions to the standard: where there is

a) specific relevant information on survival rates, age structure, birth rates and migration

or

b) clear reason to infer survival rates, age structure, birth rates and migration were within reason and in all material respects constant, proportional or irrelevant.

Survival axiom

Where survival rates are not within reason known to be constant, proportional or irrelevant, the notion of an average or other single-statistic aggregate outcome is not applicable.

Note

The standard is to cover cases where either

- a) a direct statement is made concerning aggregate outcomes or
- b) a reader might reasonably understand the claim to refer to aggregate outcomes.

An example of type (b) would be where the speaker refers to “reduction” of a condition considered undesirable and the context is of alleviating the condition for sufferers.

6. Distinguish spending from income.

Reason

In the period 1945-2005 most economic data on individuals related to spending; yet in 2005 the tradition in macroeconomics in summaries for the public was to describe the data as “income”.

Risk : The public might assume economists counted both savings and spending.

Globally, the data mostly related to spending. A minority of data were on income (notably in Latin America). Some data were on the money value of items eaten or used.

Most of the numbers were on what a sample of people said they spent.

Examples of what are strictly speaking errors:

- a) Any reference to Millennium Goal indicators as on “income”.

Millennium Goal Indicator 1 is mostly concerned with reducing the proportion of low spenders.

- b) Any reference to the economic data in policy assessments, as “income”.

For example, even if there were no other problems, a claim by economists that “poor people’s incomes rose at the same rate as policy X” would still be misleading. More accurate would be “poor people’s spending rose with policy X”.

Terminological issue: The fact that the international data are a mixture poses a linguistic problem for researchers. The problem is how to describe the data in a way both accurate and concise.

“Income” is misleading. It would be more accurate to describe the data as on spending.

7. Distinguish spending from consumption.

Purpose of standard

To ensure the phrase “consumption expenditure” is shortened to “spending” rather than “consumption”.

Note

The word “consumption” would perhaps most naturally be taken by a non-economist to refer to “items or services received or used”.

In this sense, consumption would be “things received”, not “money spent”. An economist who had data on spending and who then wrote “consumption has risen” would not be misleading the public as much as one who wrote “poverty has fallen” or “incomes have risen” but would still be misleading them.

The definition of what counts as consumption is in any case perhaps in some ways arbitrary.

8. Distinguish consumption from adequacy.

Example: Food consumption.

Food consumption adequacy for daily tasks depends on at least:

size,
age,
economies of scale,
workload type,
workload amount,
weather,
food balance,
food quality.

Examples of error

a) Macroeconomic claims on global poverty.

Per-person statistics used by World Bank, UN and others up until 2005 failed to take into account that the proportion of children is falling.

b) Policy assessments from macroeconomists based on per capita statistics.

Proportions of children are not constant across countries, or within countries by spending level.

In both cases inferences from consumption to adequacy were made without reference to needs.

It is difficult to see how a researcher who has not estimated needs for fuel, food, water, medicine, rent or other basic items could have estimated poverty.

Real-world puzzles potentially partially explained by this error:

i) Discrepancy between Food and Agriculture Organisation reported hunger trend and World Bank reported poverty trend for Millennium Goals.

ii) Discrepancy between protesters' and economists' views of progress of global poorest.

iii) Discrepancy between World Bank and other reports of progress on Millennium Goals.

Note: In some periods of history falls in child-adult ratios (which make cross-sectional per capita statistics, other things being equal, overestimate progress) may coincide with rises in longevity (which make cross-sectional statistics, other things being equal, underestimate progress).

Note on the concept of adequacy

Consumption adequacy is a concept rather than a scientific variable.

For example, it might be argued that people in a country where family members live further apart need more money for transport.

How much fuel is needed in a cold place to get to the same standard of living as in a hot place?

Such needs, as with many things described as needs, are matters of opinion, not science.

See below on inflation and cost of living for a related distinction.

9. Distinguish income from profit.

Reason

Risk of inferring gains without estimating needs.

Example

Inferences from “income” to “income poverty”.

Profit axiom

(Income) - (necessary outgoings) = (profit).

Parallel

Businesses. The axiom applies to a household as for a business.

For both businesses and households, the following are true:

Revenue, income or turnover are not profit.

1% more turnover does not indicate 1% more profit.

An income rise of 1% does not measure an income gain of 1%.

Note

It is not clear what philosophical argument might be advanced that income is an indicator of welfare.

It does not measure the cost of rent, childcare, transport, fuel, food, water or medical services in any country. It is a measure of money going round the system.

Income is a social indicator.

10. Distinguish prices from relevant prices.

Reason for standard:

Common assumption in macroeconomics that adjusting for national prices is adequate to show economic benefits for the “poorest”.

Examples of error:

Use of data on the “distribution of income” to infer consumption gains to people on low incomes.

One error in this procedure is that data are usually spending, not income.

Secondly this is a theoretical distribution in a model which does not take demographic change into account (strictly speaking, confusing distribution *among people at different times* with distribution *to real people over a period*);

Thirdly, this procedure also confuses prices with costs (more sensibly, costs = prices multiplied by needs - see below).

Also, the distribution of income would not measure the distribution of price changes.

11. Distinguish prices from cost of living.

Cost of living = prices x needs.

Common confusion in macroeconomics. Cost of living is not simply a function of (dependent on) prices.

Statements implying “statistics were adjusted for prices” to be distinguished from statements implying “cost of living has been taken into account”.

Cost-of-living axiom

The cost of living an equivalent life
in different places,
at different times,
in different circumstances and/or
for different people,
insofar as it might be assessed in some respects objectively,
necessarily depends on both

- a) prices
- and
- b) amounts needed.

Note:

It follows that inflation is not a measure of the cost of living.

Where a researcher either states or implies cost of living has been taken into account, the standard is to apply.

Note

Claims concerning needs, or the relative satisfaction of needs, are logically matters of opinion.

For example, suppose someone says people have done 1% better economically in country X than in country Y.

They might talk about income, expenses, prices of necessities and luxuries, house prices, rent prices, and so on. But in relation to human well-being, the cost of living surely means the cost of living an equivalent life.

Note on the literal cost of living

In the case of the literal cost of living - of staying alive - there is still the question of "staying alive in what condition?". This includes, for example, potential damage to brain or other organs from malnourishment.

12. Distinguish consumption gains from material gains.

The fact that I eat 1% better this year does not tell you that I am 1% better off. I might have sold my land to pay for food.

(Note: this is not meant to imply that either "I eat 1% better" or "I am 1% better" make sense).

Parallel: Business. Capital gains are relevant to both businesses and people. In both cases, to leave them out in inferring material gains is a mistake.

Capital gains and losses, including debts, are factors in a person's control or potential control of material resources. An assumption that capital gains and losses were in proportion to and in same direction as profit is merely an assumption, not a scientific finding.

In the case of people, there is the additional question of what counts as an economic gain. This question persists even if assets and other things related to legal control of resources are added in. It is not clear how a scientist who wishes to come judgements about levels of economic gains can dodge the question of what constitutes an equivalent life in different circumstances (see under cost of living, above).

13. Distinguish incidence from prevalence.

Example

Proportion of poor people cannot tell me how many were poor during period.

Prevalence (for example the prevailing number of cases now and a year ago) does not tell a researcher about incidence (the number of incidents over the year).

It is the tradition in economics to call prevalence "incidence".

The tradition among statisticians and medical researchers is to talk of the "prevalence" of hunger or disease when referring to the same concept.

Harmony of language in this respect between economics and other social sciences might be useful.

Incidence can rise and prevalence fall if more die.

14. Distinguish prevalence from degree.

The number of rich people does not say how rich they are.

15. Distinguish material conditions from judgements on well-being.

In purely physical terms, material inputs are not the same thing as bodily results.

A consumption gain, of good food to a hungry person, is not useful if they cannot absorb the food. That may depend on the state of their digestive system.

At the other extreme, many people in some countries overeat. More consumption, clearly, does not show more well-being.

In terms beyond the physical - in terms of happiness, fulfilment, and so on - it is not clear why anyone would think material conditions show overall well-being.

At the abstract level, there is a philosophical question about where a dividing line might be between “my material circumstances” and “the environment around me”. The abstract question arises because things which are legally not mine but which I use (for example trees or hills) may be functionally equivalent to legal possessions.

A reasonable line of argument might be that this does not matter very much, since other aspects of prosperity - the relative importance of life length, health, assets, debts, stability, self-respect, consumption of particular items, variety and so on - are subjective anyway.

Another line of reasoning might be this: Arguments about whether people have done “better” or “worse” under different policies are largely pointless, since some important aspects of the human condition are not possible to research.

16. Imagine real people.

One test of whether a social science claim makes sense is to think about yourself in the same position as a subject of the research.

Suppose you decide you would not apply the method to yourself. You might then reasonably decide that not to make believe the claim about someone else. This technique can be used for many types of inferences.

That test is about thinking about yourself as one person. Imagination can also be used in relation to adding up outcomes. It is often sensible to imagine a smaller unit than the research is about. For instance, suppose the claim is about a country. You can imagine a village or a street or a family. If the claim is about comparing what happens to people in different countries, you can imagine people in different houses, or streets or villages. This is really concerned with a frame of mind. If a researcher is talking about “consumption” then you are unlikely to grasp what relationship this may have to the real world unless you can imagine how it applies to one person.

Sometimes imagining people at the extremes helps bring out the basic principles.

When social scientists use abstract nouns, it is wise to see how they apply to real life.

It is a good idea not to accept what abstract nouns are supposed to be telling you unless you are clear about:

- a) what the researcher means them to refer to in real life and
- b) in what ways you think that does or does not make sense.

17. Look at meaningful groups of people.

A fundamental principle in statistics is this:

Conclusions about how one thing correlates with another are not sensible if there is no clear cut-off point for the groups.

The problem is this:

Apparently significant differences might be caused by people just inside and/or just outside the group being looked at.

If a researcher says “people who are X have a higher trend than people who are Y” it is worth asking whether the division makes sense.

18. Distinguish “statistical significance” from importance.

Statistical significance does not tell you a thing is important.

Suppose a medical researcher finds that a procedure has a statistically significant association with outcomes.

This does not tell you the procedure is significant in real life.

“Statistical significance” just means that the association seems, given the researcher’s assumptions, to be unlikely to be due to chance. It can still be a small effect, even if all the researcher’s assumptions are correct.

For example, if 61% of people taking a medicine lose their symptoms and 60% of those without the medicine lose the symptoms, the association can still be statistically significant.

So if you read a medical research paper which says a medicine is effective, an early question might be “how effective?”.

Note: Another problem which may arise in relation to claims about medical research is this: Comparing a medicine to a placebo (fake treatment) is not the same as comparing it to no treatment.

In the placebo case, according to standard practice the patient agrees to the following: they are to get something which may or may not be the real treatment.

A problem with this procedure is that the patient uses up time and energy on it (going to the doctor, remembering to take the pill or whatever it is, and thinking about whether they have got the real treatment). A “no-treatment” condition would be different from this. If the researcher sent some people away in the knowledge that they were not being given the true treatment, they would perhaps do something else.

Another problem with placebos as comparison conditions is that they are something which the patient knows may not be the real treatment. A true placebo would be one which the researcher said would work, or at least was definitely the preferred treatment.

General notes on the standards

Possible objection

“Some distinctions are not relevant, or not important in practice in discipline A or type of research B”.

Answer

In any area of work where there are reasonably foreseeable circumstances in which the distinctions may become important, it seems logical that they should be considered seriously.

Where a social scientist has information on variable x, and wishes to say something about variable y or express an opinion about z, the burden of proof is on the scientist to show the logic of the inference.

Note on types of argument

It can be useful to ask what exactly a writer is saying about economic data.

It is important for scientists - and anyone reading conclusions by scientists - to distinguish between

- a) calculation,
- b) inference,
- c) hypothesis,
- d) conjecture and
- e) opinion.

Calculation, or deduction, is where something must be true if other things are true.

Inference is where you come to a conclusion because it seems sensible.

Hypothesis is where you provide what seems a sensible explanation of the evidence.

Conjecture is a guess about either a conclusion or explanation.

Opinion, rather than knowledge, is what is possible to give about some matters. In the economic sphere, these include

- a) the relative importance of assets, debts, community assets, environmental assets, life length, culture, health, leisure time, working conditions, working hours and other aspects of economic conditions to real people;
- b) what people need.

It is also important to understand

a) where statistics come from and

b) their limitations compared to broader aspects of real life you may think important.

Context of the standards

The proposed standards are not sufficient to ensure that statements on statistics, even in the limited areas covered here, reflect real life adequately. The aim has been to identify some conditions under which descriptions of research may better reflect real life - in aspects which are likely to make significant differences to people's lives.

These proposed standards form a small subset of those which might be proposed.

One response to these proposals might be "but statistics can be misdescribed in many other ways in any case" - which is true. That would be a good starting point for further standards.

Suggested examination questions

For school courses in social science

1. What is the difference between “the average rise” and “the rise in the average“?
2. How is this difference relevant to people who are ill, elderly or poor?
3. If a statement depends on three assumptions each of 80% probability, what is the likelihood of it being right?

For philosophy courses

4. Under what circumstances would you consider the following a true statement?
“I have measured your prosperity”.
5. Under what circumstances would you consider the following a true statement?
“Poverty has fallen by more in country X than in country Y”.
6. Under what circumstances would you consider the following a true statement?
“Policy X is better for the poor than policy Y”.