B1. Methodology and Scientific Reasoning

# Simplicity as a Guide to Falsity?

Ioannis Votsis
votsis@phil-fak.uni-duesseldorf.de
University of Duesseldorf, Germany

**Short Abstract**:
Participants in the debate about whether simplicity is a guide to truth or merely pragmatically useful typically wrangle over two problems: (1) how to weigh simplicity against other virtues like strength and fitness and (2) whether there is a unique measure for simplicity that straps it to truth. I would like to put forth a third problem: (3) Even if problems (1) and (2) could be solved, it is far from clear whether the simplest theory out of an available class of competitors would always be the one closest to the truth.

**Extended Abstract:**
Participants in the debate about whether simplicity is a guide to truth or merely pragmatically useful typically wrangle over two problems: (1) how to weigh simplicity against other virtues like strength and fitness and (2) whether there is a unique measure for simplicity that straps it to truth. I would like to put forth a third problem: (3) Even if problems (1) and (2) could be solved, it is far from clear whether the simplest theory out of an available class of competitors would always be the one closest to the truth.

Hypotheses, and theories, are by their very nature simple – at least to some extent. This can be demonstrated by considering some general facts about language and classifications schemes. Most interesting hypotheses and theories in science do not concern themselves with single events, state-of-affairs, facts or whatever metaphysical building blocks one tends to favour. Rather they concern themselves with potentially recurring events, states-of-affairs, facts, etc. Otherwise put, they are (strict or statistical) regularities. Take, for example, the hypothesis that the higher the urbanised proportion of a country's population, the higher the level of its income (for a critical discussion see Bloom et. al. 2008). Regardless of the actual merits of this hypothesis, it should be obvious that it is a regularity that imposes a certain classification scheme on the world, namely populations, incomes, urban areas, etc. As such it already presupposes order and hence some degree of simplicity. Hypotheses and theories would be hopelessly complex without our ability to carve up the world into manageable chunks. For how else could we comprehend a hypothesis, communicate it to others, draw consequences from it, test it and ultimately accept, revise or reject it, if it were not to some extent simple? Whether we realise it or not, simplicity always seems to play an active role in the construction of hypotheses.

What about the question of choosing between competing hypotheses or theories. Most discussions on this matter consider simplicity to be a virtue that needs to be traded off against the virtues of strength – i.e. the magnitude of a theory's informational content – and fitness – i.e. how well that content fits the observed or measured world (see Lewis' 'best system analysis' 1999, pp. 231-236).[1] Two major problems have been identified in the literature. The first one concerns the relative weight each virtue is supposed to carry. To give an example, how do we choose between two competing theories when one has a comparative advantage in fitness while the other a comparative advantage in simplicity? The second concerns the issue of whether there exists a unique measure for simplicity that allows simplicity to function as a guide to truth.[2]

[1] In my view simplicity need not always be traded off with strength and fitness.

[2] Arguably, the notions of strength and fitness are easier to pin down than simplicity.

I would like to put forward a third problem. Suppose we encounter two or more competing theories that enjoy the same degree of strength and fitness but varying degrees of simplicity. Suppose further that the simplicity ranking between these theories remains invariant under all plausible simplicity measures. According to the received view, all else being equal, one should opt for the simplest theory. The problem with this view is that it offers no independent reason to suppose that the simplest theory is true or even that it is the closest one to the truth. It may very well be the case that the *simplest* systematisation of facts or phenomena is simpler than the true systematisation, i.e. the one that tracks the structure of the world perfectly.

A constructive demonstration of the said scenario can be made via a small detour to information theory. Compression algorithms transform data sets of a certain size into smaller data sets. Different algorithms correspond to different compression rates. Take a data set that can be compressed by various lossless algorithms.[3] Suppose that we have all these algorithms at our disposal and that we are tasked to send a compressed version of the data set to a certain receiver. Suppose further that the receiver only has access to one of those algorithms, though it is unbeknownst to us which one. Suppose finally that we only have one chance to send the compressed data set – we can imagine a third party intercepting and destroying all subsequent attempts to transmit data sets – and no possibility of feedback from the receiver. Is there a way to choose the correct – i.e. the shared – algorithm (we should compress our data set with)? Given the way I pieced together the example the answer should be 'no', or, rather, 'only by chance'. An analogous situation holds in the example of theory choice given earlier. In both examples we have a number of competing systematisations of a set: theories in the first example, compression algorithms in the second. Also in both examples there is a privileged systematisation: the theory that tracks the structure of the world perfectly in the first example, the lossless compression algorithm shared by both sender and receiver in the second. Moreover, in both examples we cannot assume, without additional information, that the privileged systematisation is the one that systematises the most: the simplest theory in the first example, the most economical compression algorithm in the second. Finally, in both examples this additional information seems to be missing: other than luck there seems to be no independent reason to suppose that the simplest theory is true or closest to the truth in the first example, other than luck there seems to be no independent reason to suppose that the algorithm we choose to compress our data is the algorithm shared by the receiver.

**References**:

Bloom, D. E. et. al. (2008) 'Urbanization and the Wealth of Nations', *Science*, vol. 319 (5864): 772-775.

Lewis, D. (1999) *Papers in Metaphysics and Epistemology*, vol. 2, Cambridge: CUP.

[3] Lossless (vs. lossy) compression algorithms allow us to recover the original data set through decompression, i.e. no content is lost in the process of compression.