To what extent do you agree with the claim "all models are wrong, but some are useful" (attributed to George Box)? Discuss with reference to mathematics and one other area of knowledge.

Django Shaw

Assessment Spring 2025

1599 words

(Excluding title page, tables, figure captions, footnotes and bibliography)

Claims addressing *all* of anything are bold, because they can be refuted by citing a single exception. Box's claim may hold regarding his field, statistics,[1] but experts in the AOKs of **Mathematics** and the **Natural Sciences** might have different criteria for 'usefulness' depending on their purposes; even the interpretations and implications of the word 'wrong' may depend on the field in question.

I start with the following definition for 'usefulness', compiled from the papers I have read which discuss models' 'usefulness': if a model gives a problem a workable solution, I consider the model useful.[2] [3] [4]

If a model ignores some affector to what it represents, I consider the model wrong, regardless of whether the model solves a problem correctly. This definition conforms with the 'no false lemmas' response to the Justified True Belief (JTB) definition of knowledge, argued by Armstrong: that a JTB based on one or several wrong assumptions, even partially, is not knowledge.[5] So, a 'right' model must replicate every affector to its target system. Sometimes, other definitions of 'wrong' will also be explored.

---

[1] "George E. P. Box (1919-2013)." 2013. The Aperiodical. https://aperiodical.com/2013/04/george-e-p-box-1919-2013/.

[2] Apostel, Leo. 1961. "Towards the Formal Study of Models in the non-Formal Sciences." In The Concept and the Role of the Model in Mathematics and Natural and Social Sciences, edited by Hans Freudenthal, 1-37: Springer Netherlands. 10.1007/978-94-010-3667-2, page 2.

[3] Mac Lane, Saunders. 1981. "Mathematical Models: A Sketch for the Philosophy of Mathematics." The American Mathematical Monthly 88 (7): 462–72. JSTOR: https://doi.org/10.2307/2321751. Accessed 4 Feb 2025, page 471.

[4] Hartmann, Stephan. 2005. "Models as a Tool for Theory Construction: Some Strategies of Preliminary Physics." (August), 1. https://philsci-archive.pitt.edu/id/eprint/2410, page 6.

[5] Armstrong, D. M. 2009. *Belief, Truth and Knowledge*: Cambridge University Press. https://doi.org/10.1017/CBO9780511570827, page 152.

The word ‘models’ is problematic: every philosopher I have consulted defines it differently.[6] [7] [8] [9] This inconsistency is a recognised issue in academia that has worsened since the 1960s.[10] [11] Most mathematical and scientific philosophers seem to share the core view that a model is a “simplified and stylised version of the so-called target system, the part or aspect of the world that we are interested in”.[12] This definition, worded by theoretical physicist Frigg, suits the purposes of this essay because it is specific to my topical AOKs but broad enough to consider a heterogenous range of models.

Within the natural sciences, I focus on physics because its similarity to mathematics means that differences between the two fields explored are inherently consequential.

This essay will explore how models’ ‘wrongness’ might be a prerequisite for ‘usefulness’, how models’ ‘usefulness’ might intrinsically make them ‘right’, and the implications of different interpretations of ‘wrongness’ on the veracity of Box’s claim.

---

[6] Giere, Ronald N. 1988. Explaining Science: A Cognitive Approach: University of Chicago Press, page 81.
[7] Frigg, Roman. 2009. “Models in Physics.” Routledge. 10.4324/9780415249126-Q135-1, page 1.
[8] (Hartmann 2005, 1)
[9] (Saunders Mac 1981, 467)
[10] (Apostel 1961, 24)
[11] (Hartmann 2005, 3)
[12] (Frigg 2009, 1)

Box implies usefulness arises despite wrongness (*"but"*). However, some abstracted models are useful *because* they are wrong.

Scientific philosopher Hartmann argues that because science is a "dynamic process" in which understandings constantly evolve, the assumption should be that current understandings are wrong.[13] In addition to being assumed wrong, representations replicating every affector to a system are cumbersome, and their accuracy may not be more useful to scientific understanding than that offered by abstracted models. Consequently, ideal models give solutions that are 'more wrong', but not so wrong as to compromise the solutions' usefulness – they must be wrong to the right extent.

Some models can be refined to various 'wrongness-levels' depending on which is most useful. In physics, where models often consist of an equation,[14] one such model is the thermodynamic 'billiard ball' model, defined by the equation $PV = nRT$. This model ignores several factors affecting thermodynamic phenomena; nevertheless, its accuracy usually suffices.[15] In situations where this model is 'too wrong' to be useful, it is refined into the model: $(p + \frac{a}{V^2})(V - b) = nRT$, sacrificing simplicity to account for both factors omitted by the 'billiard ball' model.[16] The model can be further elaborated, until decreasing 'wrongness' no longer increases 'usefulness'.

[13] (Hartmann 2005, 1)
[14] (Frigg 2009, 1)
[15] (Hartmann 2005, 4)
[16] (Hartmann 2005, 4 – The Van der Waals model.) Billiard ball model fails to account for particles' diameter or the interactive forces between them.

The MIT hadron model, from quantum physics, took this further by completely ignoring the mechanics that governed its target system.[17] It was useful because, although nuclear hadron structures were theoretically describable, computing power was insufficient to solve the equations until the 2010s. Instead, physicists used a simplified model which "contradict[ed] fundamental theory".[18] The model was only around 80% accurate, but limitations on computational power required approximations that would normally "not be sensible to make": physicists deemed wrong answers more useful than no answers at all.[19]

Physicists also use models which, according to Frigg, cannot be wrong because they do not "represent anything in the world".[20] For a budding field to properly be explored, these so-called 'toy models' are built to develop procedures on which models in the new field will be based. Toy models' idealisations "facilitate the description of an object or system", and the discoveries the models enable mean that the models "suggest their own improvements".[21]

One such model, the ϕ^4 model, describes the behavior of a subatomic particle which cannot exist. It is nonetheless useful because it "provides a frame" to describe quantum-level interaction between fields.[22] The ϕ^4 model is used pedagogically, for students to get used to the principles of quantum fields. Its defining equation is also embedded inside various other models in quantum field

[17] Johnson, K. 1975. "The M.I.T. bag model." Acta Physica Polonica: B6 (6): 865-892. https://inspirehep.net/files/7326879becf517d363853f6037b27c8e, page 885
[18] (Frigg 2009, 5)
[19] (Johnson 1975, 885)
[20] (Frigg 2009, 6)
[21] (Hartmann 2005, 4)
[22] (Hartmann 2005, 9)

theory. This is an example of a model which is useful but, in a sense, cannot be wrong.

Mathematical models tend to be less abstracted than scientific models, because pure mathematics is exact in principle.[23] Mathematics depends less on uncertainty-carrying empirical measurements than the sciences do, so mathematicians are often more expectant of exactness than scientists are.[24]

The two AOK's aims also diverge: science aims to understand, whereas mathematics aims to achieve a result – mathematician MacLane comments that, after finding an answer, the model used "does not matter and can be immediately forgotten".[25] This mindset means mathematical models' abstractions often alter real-life causality — more similarly to the MIT hadron model than to more conventional abstractional models.

MacLane expands the definition of 'models' to include simplification as well as abstraction. Through 'cross-modelisation', where a problem is represented through a different branch of mathematics, new methods of solution can be found without abstracting the target system (so 'cross-modelisation' is useful without being wrong).[26] There is doubt as to whether 'cross-modelisation' is a model, as it simplifies without abstracting, but it serves the purpose ascribed to models, so it might be considered one.

[23] Liu, Jerry Z. n.d. "Science vs. Mathematics." *Stanford University, California, USA*, 1-10. Accessed February 15, 2025. https://cs.stanford.edu/people/zjl/pdf/science.pdf. Page 1.
[24] (ibid.)
[25] (MacLane 1981, 467)
[26] (MacLane 1981, 463)

'Cross-modelisation' is possible because all mathematical branches stem from the same set of human activities (counting, measuring, navigating…) which were refined and combined into progressively more complex rulesets (*E.g. Fig. 1)*. Put more classically, *all areas of mathematics are based on the same set of axioms*.[27]

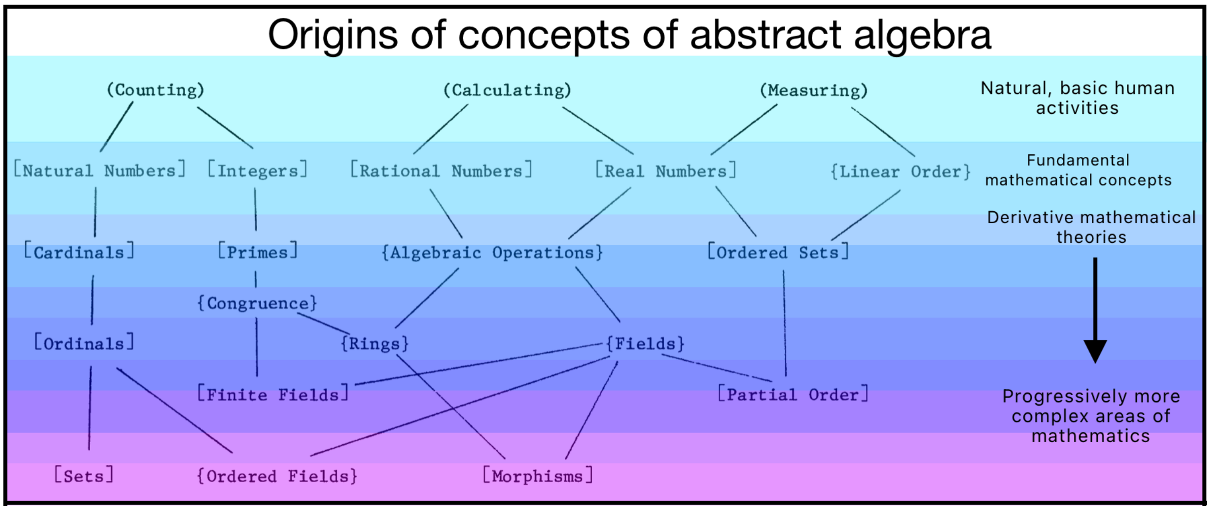

Fig.1 -- Example illustration of the web of mathematical concepts derived from basic human activities, combining and feeding off each other, ending (at the bottom) with abstract algebraic concepts. (Original web from Saunders Mac 1981, 464. *Edited by me.)*

One example of cross-modelisation is assigning coordinates to a euclidean space, to turn geometric problems into algebraic ones. In my mathematics IA, I used cross-modelisation to solve a geometric problem involving the distance to a mountain. I derived the following equation, which algebraically models the three-dimensional environment (*next page*):

[27] Argued by David Hilbert in *Grundlagen der Geometrie (Festschrift 1899)*: Springer Berlin Heidelberg.

$$\delta = \frac{\pi r}{180}\cos^{-1}\left(\frac{\sqrt{\dfrac{\Delta h^2\left(\dfrac{\cos\omega_1+1}{\cos\omega_2+1}\cdot\dfrac{\sin\omega_2}{\sin\omega_1}\right)^2}{1-\left(\dfrac{\cos\omega_1+1}{\cos\omega_2+1}\cdot\dfrac{\sin\omega_2}{\sin\omega_1}\right)^2}}}{r+h_p}\right)$$

(Where ∂ represents distance to the mountain)

There is debate as to whether *every* area of mathematics stems from the same axioms – Gödel proved that for any given set of axioms, there exists a claim that cannot be proven or disproved using these axioms.[28] Nevertheless, 'cross-modelisation' was useful to me, and cannot be wrong because it involves no abstraction.

Having considered the usefulness and wrongness of models outside of statistics, Box's area of expertise should be examined.[29]

As a mathematical field, statistics is unusual: it concerns itself primarily with requirements of industries traditionally considered to be 'outside mathematics'. These industries, which include government, transport, and health, often prioritise statistics' usefulness over exactness. To accommodate this, statistical models consist of considerably idealised assumptions (statistician Cox).[30]

[28] Raatikainen, Panu. 2018. "Gödel's Incompleteness Theorems." Edited by Edward N. Zalta. *The Stanford Encyclopedia of Philosophy*, (Fall). https://plato.stanford.edu/archives/fall2018/entries/goedel-incompleteness/.
[29] ("George E. P. Box (1919-2013)" 2013)
[30] Cox, D. R. 2006. Principles of Statistical Inference. N.p.: Cambridge University Press, page 178.

Perhaps the simplest statistical model is the assumption that a flipped coin will show ‘heads’ 50% of the time.[31] [32] This model upholds Box’s claim: it is useful despite being wrong.

With two coin-flips, the model's prediction of a 50/50 spread is wrong in half of all sets of results (*Fig. 2*).

Larger sample sizes make the model *almost right* more often, but *exactly right* less often. Furthermore, odd numbers of flips will *never* show 50% heads.

Fig. 2 – Representation of the outcomes of the model: ‘a coin toss will result in “heads” 50% of the time’, for two flips.

1st toss	2nd toss	Model correct?
Heads	Heads	NO (100/0)
Heads	Tails	YES (50/50)
Tails	Heads	YES (50/50)
Tails	Heads	NO (0/100)

The errors inherent in estimations are magnified, generating inaccuracy. Statistical models still serve their purposes of giving reasonable estimations of particular events’ likelihood of happening. They are useful *despite* usually being wrong.

Pragmatists interpret ‘truth’ as a quality of a statement that has been challenged rigorously but not yet successfully – Richard Feynmann writes “we never are definitely right, we can only be sure we are wrong”.[33] It follows that models are ‘right’ if they provide answers of an “indicated type and degree of similarity” to the

[31] Stang, Andreas, and Bernd Kowall. 2020. “Fishers Signifikanztest: Eine sanfte Einführung.” GMS Medizinische Informatik, Biometrie und Epidemiologie, (May). 10.3205/mibe000206. Page 1
[32] Heumann, Christian, Helmut Küchenhof, and Shalabh Shalabh. 2008. “Coin Tossing and Spinning – Useful Classroom Experiments for Teaching Statistics.” Recent Advances in Linear Models and Related Areas, (July), 417-426. 10.1007/978-3-7908-2064-5_23. Page 417.
[33] Feynman, Richard P. 1994. The Character of Physical Law. N.p.: Modern Library, page 158.

(theoretical) absolute correct answer.[34] Functionally, this definition resembles my introductory definition for usefulness. Therefore, if a model's 'correctness' is defined by the pragmatic theory of truth, then being useful is its only prerequisite for correctness.

However, Goodman argues that the concept of similarity is too vague for science.[35] Other philosophers (Giere, Churchland) differ: because humans think "based on some sort of similarity metric", extending the precision of science beyond human understanding is pointless.[36] [37] Like pragmatists, subscribers to this 'correspondence theory of truth' consider an abstracted model 'right' if it represents its target system with an adequate degree of similarity.[38] Having made his point, Giere goes so far as to dismiss the debate as pointless: "A 'theory of truth' is not a prerequisite for an adequate theory of science".[39]

Considering the ideas explored above, I disagree to a large extent with the stimulus. While Box's claim has been deemed valid within his field, it does not hold when considering models outside statistics: not all models are wrong, and a model's wrongness (or lack thereof) is not a reliable indicator for its usefulness. I disagree especially with Box's implied causality (that models are useful *despite*

---

[34] (Giere 1988, 93)
[35] Goodman, Nelson. 1973. "Seven Strictures on Similarity." In Problems and Projects, 437-446. 1st Edition ed. New York/Indianapolis: Hackett Publishing Company, Incorporated. https://gwern.net/doc/philosophy/ontology/1972-goodman.pdf. Page 437.
[36] (Giere 1988, 81). Quote is from here.
[37] Churchland, Patricia S. 1989. Neurophilosophy: Toward a Unified Science of the Mind-Brain: BRADFORD BOOK. Page 456.
[38] David, Marian, and Edward N. Zalta (ed.). 2022. "The Correspondence Theory of Truth." The Stanford Encyclopedia of Philosophy. https://plato.stanford.edu/archives/sum2022/entries/truth-correspondence/.
[39] (Giere 1988, 93)

being wrong), because those models that *are* wrong and useful tend to be useful *because* they are wrong. The billiard ball model was shown as an example of this, and the MIT hadron model was presented as a replacement for a model which was useless *because* it was right.

It is important for mathematicians and scientists to consider the types of models they are using. The assumptions that the model is based on, and whether the model's status as "wrong" or "right" has any bearing on the scope and extent of the model's usefulness, are important factors to consider before coming to conclusions.

Fig. 3 shows the types of models discussed in this essay.

Fig. 3 – Table of types of models discussed in this essay.

Model name	**Wrong?**	**Useful?**	**Causality**
MIT hadron model	YES	YES	Useful because wrong
Billiard ball gas model	YES	YES	Useful because wrong
None discussed (trivial)	YES	NO	Models that are wrong and useless
ϕ^4 model	NO	YES	Cannot be wrong; useful
'Cross-modelisation'	NO	YES	No abstraction, still useful
Hadron model that the MIT hadron model replaced.	NO	NO	Correct (no abstraction), still useless

Works Cited

Apostel, Leo. 1961. "Towards the Formal Study of Models in the non-Formal Sciences." In *The Concept and the Role of the Model in Mathematics and Natural and Social Sciences*, edited by Hans Freudenthal, 1-37: Springer Netherlands. 10.1007/978-94-010-3667-2.

Armstrong, D. M. 2009. *Belief, Truth and Knowledge*. N.p.: Cambridge University Press. https://doi.org/10.1017/CBO9780511570827.

Churchland, Patricia S. 1989. *Neurophilosophy: Toward a Unified Science of the Mind-Brain*. N.p.: BRADFORD BOOK.

Cox, D. R. 2006. *Principles of Statistical Inference*. 1-236: Cambridge University Press.

David, Marian, and Edward N. Zalta (ed.). 2022. "The Correspondence Theory of Truth." The Stanford Encyclopedia of Philosophy. https://plato.stanford.edu/archives/sum2022/entries/truth-correspondence/.

Feynman, Richard P. 1994. *The Character of Physical Law*. 1-192: Modern Library.

Frigg, Roman. 2009. "Models in Physics." *Routledge*. 10.4324/9780415249126-Q135-1.

"George E. P. Box (1919-2013)." 2013. The Aperiodical. https://aperiodical.com/2013/04/george-e-p-box-1919-2013/.

Giere, Ronald N. 1988. *Explaining Science: A Cognitive Approach*. 1-344: University of Chicago Press.

Goodman, Nelson. 1973. “Seven Strictures on Similarity.” In *Problems and Projects*, 437-446. 1st Edition ed. New York/Indianapolis: Hackett Publishing Company, Incorporated. https://gwern.net/doc/philosophy/ontology/1972-goodman.pdf.

Hartmann, Stephan. 2005. “Models as a Tool for Theory Construction: Some Strategies of Preliminary Physics.” (August), 1. https://philsci-archive.pitt.edu/id/eprint/2410.

Heumann, Christian, Helmut Küchenhof, and Shalabh Shalabh. 2008. “Coin Tossing and Spinning – Useful Classroom Experiments for Teaching Statistics.” *Recent Advances in Linear Models and Related Areas*, (July), 417-426. 10.1007/978-3-7908-2064-5_23.

Johnson, K. 1975. “The M.I.T. bag model.” *Acta Physica Polonica:* B6 (6): 865-892. https://inspirehep.net/files/7326879becf517d363853f6037b27c8e.

Liu, Jerry Z. n.d. “Science vs. Mathematics.” *Stanford University, California, USA*, 1-10. Accessed February 15, 2025. https://cs.stanford.edu/people/zjl/pdf/science.pdf.

Mac Lane, Saunders. 1981. “Mathematical Models: A Sketch for the Philosophy of Mathematics.” *The American Mathematical Monthly* 88 (7): 462–72. JSTOR: https://doi.org/10.2307/2321751. Accessed 4 Feb 2025.

Raatikainen, Panu. 2018. “Gödel's Incompleteness Theorems.” Edited by Edward N. Zalta. *The Stanford Encyclopedia of Philosophy*, (Fall). https://plato.stanford.edu/archives/fall2018/entries/goedel-incompleteness/.

Stang, Andreas, and Bernd Kowall. 2020. “Fishers Signifikanztest: Eine sanfte Einführung.” *GMS Medizinische Informatik, Biometrie und Epidemiologie*, (May). https://dx.doi.org/10.3205/mibe000206.

Volkert, Klaus, ed. 2015. *David Hilbert: Grundlagen der Geometrie (Festschrift 1899)*: Springer Berlin Heidelberg.