***A Subjective Bayesian Response to Winsberg’s use of the ‘Adequacy-for-Purpose’ Model Criterion***

John Lepp

ABSTRACT: It will be argued that Eric Winsberg has created a problem where nobody is in the position to rationally support the Anthropogenic Climate Change hypothesis, since he demands the normal lay public defer to experts but, from Winsberg’s philosophical commitments, experts are precluded from having the ability to rationally conclude that a hypothesis is superior to an alternative. Winsberg’s difficulties can be resolved with a little help from Bayesian Confirmation Theory. A Bayesian analysis will be provided which allows for the evaluation of climate models that uses Winsberg’s analysis of robustness.

 Introduction

The Anthropogenic Climate Change Hypothesis - ACCH - is the hypothesis that human activity contributes to climate change. The data which supports the hypothesis “Human activity contributes to climate change” is incomplete. There are a sundry of issues in establishing a causal relation between human activity and climate (2008,3: 2022,54):

How reliable is the global temperature record?

What is the average precipitation rate both globally and for a locality?

How should clouds be treated?

What is the effect of Carbon emissions over time?

The list of issues is not meant to be understood as either sufficient or complete. With so many unresolved issues, it is difficult to establish causal agency.

In light of there being so many unknowns, the hypothesis “Human activity contributes to climate change” is underdetermined. Other explanations for climate change are, *prima facie*, possible. Other explanations have been offered. One such alternative has been offered by PhD scientists sponsored by the Heartland Institute (a conservative think tank.)

In their various documents, the Scientists from the Heartland Institute attempt to give the issues just raised a scandalous glow. The scientists associated with the Heartland Institute contend that the evidence shows - contrary to the belief that mankind has brought about harmful climate change - there is little evidence for anthropogenic climate change, and the changes in temperature and carbon dioxide concentrations in the atmosphere are beneficial for humans, plants and wildlife.

 The Heartland Institute, under the guise of ‘the Nongovernmental International Panel on Climate Change’ (NIPCC), responds to the United Nations’s ‘Intergovernmental Panel on Climate Change’ (IPCC) position paper. NIPCC summarizes their position thus:

“On the most important issue, the IPCC’s claim that ‘most of the observed increase in global average temperatures since the mid-20th century is very likely (defined by the IPCC as between 90 to 99 percent certain) due to the observed increase in anthropogenic greenhouse gas concentrations,’ NIPCC reaches the opposite conclusion – namely, that natural causes are very likely to be the dominant cause” (2008,IV).

The Heartland Institute puts forth the hypothesis, HI, that climate change is primarily a natural occurrence and there is little evidence that human activity is the main driver for climate change. They do not say that human activity cannot affect climate, but, rather, the significance of human activity is greatly overstated. The scientists from the Heartland Institute do not accept ACCH. That said, HI does not contradict ACCH.

Considering climate models are underdetermined, a means to rationally favor one hypothesis over the other would be beneficial. Grant, if only for the sake of argument, that there are several models that contain ACCH and at least one model (sponsored by the Heartland Institute) which contains an hypothesis HI denying Anthropogenic Climate Change as a main driver of climate change. What makes one hypothesis superior to another?

Eric Winsberg has recently published a very impressive tome “Philosophy and Climate Science” examining a variety of epistemological issues concerning climate models. Winsberg has indirectly answered why ACCH is preferable to the hypothesis offered by the Heartland Institute, but Winsberg's answer, given his philosophical commitments, is unsatisfactory. His answer is to note that “Climate science is scrutinized by a highly skeptical and very well-funded opposition” (2018, 215) and to insist the lay public defer their assessments of Climate Science to the Scientific community. It will be shown why this answer is unsatisfactory. A means to show the superiority of ACCH over HI will be provided.

The basic argument for this essay is as follows:

Winsberg’s adoption of the “adequacy-for-purpose" criterion for models precludes him from having the ability to rationally decide which hypothesis is better, that is to say which hypothesis is more likely true. Hence the use of “adequacy-for-purpose” precludes Winsberg from his ability to show the superiority of the ACCH over HI. Ironically, Winsberg also provides a metric which could be used to show superiority. Winsberg could show ACCH is superior to HI by using his analysis of “robustness,” but, were he to do so, he would have to abandon “Adequacy for Purpose” and allow himself to become Bayesian.

It will first be explained why Winsberg’s answer is unsatisfactory. Two concepts employed by Winsberg - “Models” and “Adequacy-for-purpose'' - will be briefly examined. Those concepts will be used to compare ACCH and HI. The Winsberg problem will then be elucidated. Subsequent to elucidating the Winsberg problem, we will begin to show how Winsberg’s problem can be resolved. The concept of “robustness” will be analyzed. “Robustness” will provide a virtue by which to establish the superiority of one hypothesis over another.

 Models, hypotheses, truth and adequacy for purpose

Winsberg adopts a conception of “model” (or data model) from Patrick Suppes. In Suppes’s analysis, an hypothesis is not tested directly against experimental results, but, rather, an hypothesis is selected by comparing what a hypothesis states with what data is put forth in a model. Models of data organize, idealize, coordinate raw data of experience to form something coherent (1962). Hypotheses are formed and chosen from those rectified results (1993, 126).

Models contain idealized assumptions and various fictions. Models are structures of the relations of data. Structures do not appear to be anything factual. They are too abstract. In light of so much that is known to be false in models, and in light of the abstract nature of models, how can models be verified to be true? Wendy Parker suggests, and Winsberg endorses, rather than test to see if a model is true, models are judged by whether or not they are adequate for purposes. In other words, Parker and Winsberg are suggesting the way to evaluate a model is to see whether or not it services meeting a goal.

Parker explains the adequacy-for-purpose criterion for models in this way; “To ask whether a model is adequate for a purpose is to ask something like: Can the model be used to do the job?” (Parker 2020, 461) What counts as “success” for a model depends on whatever is the goal for (whatever is the job to be performed by) a model. Now, as mentioned prior, attaining the truth would not be a goal for models.

The role of hypotheses is to provide guidance on how to apply models to the actual world. Hypotheses are chosen by scientists for their models. Given this role supplied by Suppes and others, an hypothesis would not be judged as “true” or “false” according to whether a hypothesis corresponds to the actual world, but, rather, by the adequacy of a model's application to the real world.

Parker provides a metric for how to evaluate the application of a model to its target, the real world:

“When considering model-target fit, the adequacy-for-purpose evaluator will focus on aspects (and degrees of fit) that are considered most relevant for achieving the purpose of interest, whereas the evaluator of representational accuracy might well employ general or overall measures of model-target fit” (2020, 471).

“The purpose of interest” is not something already set. The account of “achieving the purpose of interest” (ie success) is relative to the interests of an evaluator.

Apply an “adequacy for purpose” approach for evaluating the hypotheses ACCH and HI. Clearly the models that include ACCH

would have different goals than the model with HI. Observe that even were we to suppose that the concern of the scientists from the Heartland Institute was entirely monetary, say the goal is to explain how it would be possible to account for the raw data with the least amount of economic expense, it is not as though other scientists could have superior goals. It would all be relative to the desires of what a Scientist wants to achieve.

Return to Winsberg’s observation that “Climate science is scrutinized by a highly skeptical and very well-funded opposition.” It is hard to see why that matters. Surely Winsberg does not believe the well-funded opposition is deceiving the public into believing something which is not true. Upon replacing “truth” as a goal with whether or not a model is adequate for purpose, it is unclear what the public gains from deferring to the judgment of a less critical Scientific community. That “Climate science is scrutinized by a highly skeptical and very well-funded opposition” is too weak a reason for the lay public to either accept or reject climate science , given that no position is more likely true than any other.

A Problem, a Solution, a Revised Problem and a Revised Solution

Winsberg believes we are all faced with a problem of underdetermination due to the need of employing an ensemble of different models with different purposes. He argues that the problem of underdetermination faced by scientists dealing with a variety of models (inter-model holism) is essentially different from the problem faced by scientists dealing with the success or failure of predictions (confirmation holism.) Perhaps there is a difference, but the problem of underdetermination for inter-model holism can be resolved in much the same way as how the problem of underdetermination is resolved for confirmation holism.

The argument for Inter-model holism is as follows (2018, 139):

Multiple models are accepted by Scientists due to those models being “adequate-for-purpose,” not due their being “true” nor for their “overall fit.”

Models are constructed with purposes in mind.

Different models have different concerns.

Therefore

Mutatis mutandis, different models will have different criteria by which to judge “success.”

Denote “the syllogism for inter-model holism” as IM.

We have (when examining ACCH and HI) two competing hypotheses that are in models which were constructed with different purposes in mind. One hypothesis demands drastic change. The other hypothesis does not demand anything from us, but, if it is false, it surely means our demise. Hence we are here having to decide whether we face an existential threat. Those models were constructed with different purposes, yet it still appears we face a conflict.

Before attempting to unwind the tangled web we have found ourselves in, let us take a quick glance at “confirmatory holism” and how a Bayesian would resolve the Duhem-Quine paradox. Willard van Orman Quine strengthened an insight put forth by Pierre Duhem. Duhem’s insight was that evidence never directly refutes a hypothesis (1954, 187.) Rather, the evidence conflicts with a hypothesis and at least one other auxiliary statement. Auxiliary statements state background assumptions (which would be conditions allowing for the refutation of an hypothesis.) The Duhem-Quine paradox is as follows: Upon receiving a recalcitrant result, how do we determine what is at fault? Bayesian probability has been used to resolve this problem .

Review the following to become acquainted with the basics of Bayesian probability:

P(h) is the probability P of hypothesis h before any evidence has come into consideration. P(h) represents the prior degree of belief in h (i.e. the “prior probability” of h) before any additional evidence has come to light.

P(e) is the model evidence.

P(e/h) is the likelihood of the evidence e given the hypothesis h.

P(h/e) is the probability P of hypothesis h given in light of evidence e. P(h/e) is the posterior probability.

P(h/e) = $(P(h)P(e/h))/P(e)$ is Bayes’ theorem for Posterior probability.

 The Duhem-Quine paradox is a problem about a supposed inability to apportion praise or blame on an individual hypothesis from auxiliary statements. Once we are able to apportion praise or blame for successes or failures, the paradox will be solved.

Michael Strevens provided the standard Bayesian solution (2001, 516) to the Duhem-Quine paradox. Examine the effect that the negation of “s” (an auxiliary statement) has on our ability to determine the probability of “h.” Let “\*” be the operator for a posterior probability, so that

 P\*(h) = P (h/e)

= P (hs/e) + P (h-s/e)

= P (h-s/e) [Given that we are examining the negation of s.]

= (P(h-s) P (e/h-s)) / P(e) [See 5]

= (P(h-s) P(e/h-s))/ P(h-s) P(e)/ P(h-s) + P(hs) P(e) + P(-hs)P(e) + P(-h-s) P (e/ (-h-s)) [see 9 and the definition of a tautology.]

1- 10 provides a Bayesian means to resolve the problem of underdetermination for confirmatory holism.

Winsberg has challenged the efficacy and ability to use Bayes theorem to siphon off probabilities for hypotheses from other valuative considerations, other cultural considerations and/or psychological considerations (2012, 129; 2018, 145.) He noted, and this will not be disputed, that different models have different concerns and therefore different criteria for success. He also noted, and this too will not be disputed, that value judgements are necessarily employed in scientific practice (2018,147.)

What will be disputed here is whether or not this different criteria necessarily hampers our ability to emphasize or deemphasize criteria (or normative issues, or value judgements) so as to allow examination of hypotheses. Can Scientists consider hypotheses apart from the models that detect them? Are scientists always unable to quell their passions for models?

It should be noted that models, not hypotheses and not ensembles, are mentioned in IM. We are attempting to solve a problem of competing hypotheses (ACCH and HI.) Hypotheses are how models are to be applied to the actual world.

The standard Bayesian solution for the probability of hypotheses can be amended to handle not only auxiliary statements (some of which might have truth values) but also purposes, virtues, and other dispositions (which do not have truth values.) Hypotheses can be isolated from whatever else is contained in models using Bayesian calculus. Once hypotheses are isolated, Winsberg's problem of underdetermination from Inter-model holism will be solved.

We should amend s from denoting “a auxiliary statement” to “the set of auxiliary statements, intended purposes held, virtues (values) applied and/or dispositions held, and whatever else is contained in a model.” Admittedly, s would be a motley set and the practice of using the standard Bayesian solution would be a little too cumbersome for practicing scientists but all that was meant to be shown is that it is possible to isolate the probability of a hypothesis from auxiliary statements, and intended purposes. No need for a muddy mess where scientists are always unable to separate probabilities from their passions or from their assumptions.

4. The Virtue of Robustness

Parker attempts to refute the idea that agreement between models on hypotheses provides reason for increased confidence in the truth of such hypotheses in Climate Science. She criticizes Bayesian reasoning. Parker provides a succinct summary of an example of such Bayesian reasoning (2011, 590.):

“1. e warrants significantly increased confidence in predictive hypothesis H if p(e/H) >> p(e/-H)

 2. e = all of the models in this ensemble indicate H to be true.

3. The observed agreement among models is substantially more probable if H is true than if H is false; that is p(e/H) >> p(e/-H)

∴ e warrants significantly increased confidence in H”

Parker, however, objects that besides the data being prone to error and idealization, the models are borrowing from each other in regards to the hypotheses held, and thus concludes;

“Perhaps with additional reflection and analysis, persuasive arguments for p(e/H) >>p(e/-H) can be developed in some cases, but at present such arguments are not readily available.”

Were agreement between models the only reason for increased confidence in hypotheses, were agreement between models the only signal of robustness for hypotheses, the conversation would focus on whether Parker’s argument is persuasive. No need, though. Winsberg provides an additional examination of Robustness Analysis (RA.) Three questions arise:

Can Winsberg’s revised conception of Robustness be adapted to work with a similar syllogism as that which was provided by Parker? Related to question 1,

Need “e = all of the models in this ensemble indicate H to be true” be the case in order for the Bayesian syllogism to work? Can e equal something else?

Does RA shed any light concerning the conflict between ACCH and HI?

Winsberg’s analysis shows the viability of Bayesian reasoning, and light will be shed on the conflict between ACCH and HI using his analysis.

Winsberg wholeheartedly agrees with Parker’s replacing the notion of truth with adequacy for purpose,

“Parker is right that climate models are best discussed in terms of their adequacy for purpose, rather than their truth, or overall fit. At the very least, therefore, we are going to have to replace the notion of truth in this picture of RA with the idea of adequacy for purpose” (2021, 55103).

Remember to replace “true” with “adequate for purpose” each time the word “true” is mentioned and see if what you are reading still makes sense.

Winsberg parts ways from Parker by examining the set of procedures (which can be from models, but can also be from deductions, experiments, or observations) which indicate that a hypothesis is true. Winsberg develops an analysis of “Robustness Analysis” - based on Joel Schupbach’s work (2016b) - where robustness for a hypothesis is measured against a rival hypothesis. Winsberg introduces the metric ‘Cumulative Epistemic Power’ (CEP). As the set of procedures which indicate the truth of a hypothesis (call that set of procedures D) grows, so grows D’s robustness, so grows the CEP of a hypothesis.

It is difficult, if not impossible, to measure just how large the set of examining procedures should be isolated from other hypotheses. A rival hypothesis establishes a basis by which to judge preferability of set size. Simply choose the larger set.

Winsberg, following Schupbach, would have the addition of each procedure added to D eliminate a competing hypothesis (S5112). That requirement is a little too restrictive to allow for the best usage of CEP. Rival hypotheses sometimes directly contradict hypotheses, but not always. Sometimes the truth of other hypotheses is granted. The requirement provided by Schupbach and Winsberg can be adjusted to allow for CEP without altering the force of CEP.

The following example sheds light on the problem with Schupbach and Winsberg's requirement:

Have H be understood to be the hypothesis “Water freezes at 0 degrees celsius.” Let -H be its negation “Water does not freeze at 0 degrees celsius.” Let H’ be the hypothesis “Warlocks freeze Water as desired.” H has, at least, two other competing rivals. Upon sticking a plastic cup of water in a refrigerator set at 0 degrees celsius for a day, H will be confirmed, -H will be disconfirmed, and H will be better confirmed than H’ since H is more likely true.

Let us further suppose that supporters of H’ do not deny that water freezes at 0 degrees celsius. It is just that they believe that water freezing is predominantly due to the desires of warlocks. Even in this instance, we would still be inclined to believe that the evidence of water freezing better confirms H than it does H’. H is more likely true. It just does not matter whether supporters of H’ accept the possibility of H or not. That said, confirming instances of H do not eliminate H’.

Of course, the set of procedures that indicate H is more robust than H’. There are a variety of different procedures which indicate H is true. H is made evident from placing a cup of water in the refrigerator at 0 degree celsius, observing water in a lake freeze when it is 0 degree celsius, as well as by understanding Chemistry and the low energy state of molecules at 0 degree Celsius, resulting in the open lattice of hydrogen bonded molecules that are to be found at 0 degrees Celsius. There are pretty much no procedures which indicate anything about warlocks.

The CEP for H is greater than H’. The set of procedures indicating the truth of H is greater than the set of procedures indicating the truth of H’. The point of the warlock example is to show that Schupbach’s and Winsberg’s requirement can be made to be less restrictive without stopping the notion of CEP from doing its work.

Winsberg’s CEP analysis works well with a revised Bayesian syllogism:

1. e warrants significantly increased confidence in predictive hypothesis

H if p(e/H)>> p(e/-H) and p(e/H) >> p(e/H’)

2. e = whatever the method of detection finds that indicates H is more likely true than -H or H’.

3. The detections are substantially more probable if H is true than if H is false; that is p(e/H)>> p(e/-H) Likewise, the detections are substantially more probable should H be true than what would be the probability of detections were H’ true . That is p(e/H)>> p(e/H’).

∴ e warrants significantly increased confidence in H.

5. Robustness and Our Competing Hypotheses

“Natural causes” are here understood to be that which occurred not due to mankind’s behavior. The negation of “natural causes” is to be understood as anthropogenic. Examine the following evidence for ACCH:

A)

 $ Carbon emissions from 1850 to 2011 have risen 1500% <https://www.wri.org/insights/history-carbon-dioxide-emissions>.

 $$ A long term warming trend for the Earth’s temperature (both land and sea) has been measured for the dates from 1880 to 2014. <https://www.climate.gov/news-features/videos/history-earths-temperature-1880>

The two diverse data points point to an increase in the likelihood of the truth of ACCH. There is an apparent correlation. The apparent correlation does not itself prove a causal relationship between $ and $$, but it does raise its likelihood. Other data points are needed to increase the probability of the truth for such a relationship.

There are other detection procedures pointing towards the truth of ACCH.

B)

@ The troposphere is warming.

@@ The rest of the atmosphere is not warming. <https://news.climate.columbia.edu/2017/04/04/how-we-know-climate-change-is-not-natural/>

Warming from the sun, a natural cause, would be uniform. The wrong result. It is an anomaly what other natural causes would bring about the uneven warming. The warming of the troposphere, however, is the expected result were the warming due to man made carbon emissions.

C)

Between 1975 to 2020, ocean temperatures have been rising faster than they would by known natural causes. Anthropogenic causes would explain the rise.

<https://www.epa.gov/climate-indicators/climate-change-indicators-ocean-heat>

A, B, and C are detection procedures for set D. There are other procedures but there is no need to enumerate them all here.

Denote the set of detection procedures for HI by N.

There are detection procedures to be added to N, but the procedures are of a far more limited nature. For instance, the observation “Solar activity can affect climate” can be added to N. The CEP of the procedures added to N is weaker than the CEP of procedures added to D.

The scientists from the Heartland Institute are not trying to make HI more robust by explaining ever more diverse evidence. They are, rather, attempting to maintain the strength of HI by denying evidence for its rival. Confidence in HI survives by maintaining the status quo on accepted evidence versus a rival.

Notice, for instance, that the model containing HI did not supply an alternative explanation for A. It attacks the data itself (2016, 69.) The Heartland Institute called into question $$ that temperatures have been rising. For B, while the Heartland Institute accepts the troposphere is warming faster than the rest of the atmosphere, they deny the data supports the effect of carbon emissions (2016.) For HI, they do not provide an alternative explanation for C, they simply reject C (2019).

The strategy for supporting HI is pretty much the inverse of CEP for ACCH. The strategy for supporting HI is Subtracting Epistemic Power SEP from ACCH. Perhaps HI gains a modicum of strength if ACCH has less CEP, but is this a preferable way to establish an hypothesis? It would seem, upon a little reflection, that it could only lead us to become more agnostic as to the truth of an hypothesis but nothing more. The force of Winsberg’s observation that “Climate science is scrutinized by a highly skeptical and very well-funded opposition” increases when it is understood that the well-funded opposition obfuscates the truth of hypotheses.

CEP provides robustness. What does SEP provide? A shrugging of shoulders? Indifference? It is difficult to find a virtue to associate with SEP. With a consistent application of SEP, the best result for HI would be for HI to be seen as robust as ACCH. Above that baseline, SEP does not strengthen HI. After a certain point, HI would not be found to be more robust from the application of SEP.

 As we saw with the hypothesis about warlocks, it simply doesn't matter if the Heartland institute also grants the possibility of anthropogenic causes. They need positive evidence to support their hypothesis. Their hypothesis HI is not as robust as ACCH.

Back to Winsberg’s problem of evaluating hypotheses. Can experts tell if ACCH is superior to HI? Yes they can. We can. ACCH is more robust, and that provides good reason to consider it more likely true.

6. Conclusion

Having shown the error of Winsberg's ways, should you wish to allow for either the lay public to follow expert opinion or for expert opinion to be justified, you can at least respond in one of two ways to the proposed amendments for Winsberg’s work made here:

You can accept the Bayesian amendments here proposed so as to allow justified expert opinion about the veracity of the anthropogenic hypothesis; or

 You can continue to accept Winsberg's argument without revision, and just choose to remain irrational.

It is, of course, your choice.

References

Duhem, Pierre (1954). *The Aim and Structure of Physical Theory* (Translation by Phillip Wiener of La the’orie physique son objet et sa structure, 2nd ed. Paris, Chevalier et Riviere, 1914). Princeton: Princeton University Press.

 Idso, Craig D.Ph.D., Legates, David Ph.D., Singer, S. Fred Ph.D. “Climate Change Reconsidered” Chapter 2 p.19 (2009) *Report of the Nongovernmental International Panel on Climate Change* (NIPCC) The Heartland Institute

Idso, Craig D.Ph.D., Singer, S. Fred Ph.D. “Why Scientists Disagree About Climate Change” (2016) *Report of the Nongovernmental International Panel on Climate Change* (NIPCC) The Heartland Institute

Intergovernmental Panel on Climate Change *Climate Change 2022: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change* (2022) edited by H.-O. Pörtner, D.C. Roberts, M. Tignor, E.S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, B. Rama Cambridge, UK and New York, NY, USA: Cambridge University Press. 3056 doi:10.1017/9781009325844

Parker , Wendy “When Climate Models Agree: The Significance of Robust Model Predictions” (2011) *Philosophy of Science*, Vol. 78, No. 4: 579-600

Parker, W & Winsberg, E “Values and Evidence: How Models Make a Difference” (2017) *European Journal for the Philosophy of Science* Vol.8: 125-142

 Strevens, Michael “The Bayesian Treatment of Auxiliary Hypotheses” *British Journal for the Philosophy of Science* 52 (2001): pp.515-537

Quine, W.V.O. “Two Dogmas of Empiricism” *From a Logical Point of View* (1953) Harper Torchbooks: New York

Schupbach, Joel (2016b). Robustness analysis as explanatory reasoning. British Journal for the Philosophy of Science. https://doi.org/10.1093/bjps/axw008.

Suppes, Patrick “Models of Data” (1962) E. Nagel, P. Suppes, & A. Tarski (Eds.), *Logic, Methodology, and Philosophy of Science: Proceedings of the 1960 International Congress* (1962) Stanford: Stanford University Press, 252-261.

Suppes, Patrick “Models and Methods in the Philosophy of Science: Selected Essays” (1993) J. Hintikka Editor Synthese Library, Vol 226

Whitehouse, David “Cold Water? The Oceans and Climate Change” (2019) https://www.heartland.org/publications-resources/publications/cold-water-the-oceans-and-climate-change

Winsberg, Eric “Values and Uncertainties in the Predictions of Global Climate Models” *Kennedy Institute of Ethics Journal* (2012) Volume 22, Number 2, 111-137

Winsberg, Eric *Philosophy and Climate Science* (2018) New York: Cambridge University Press.

Winsberg, Eric (2021) “What does robustness teach us in climate science: a re-appraisal” Synthese 198 (Suppl 21):S5099–S5122 https://doi.org/10.1007/s11229-018-01997-7t