IOPS Winter 2021

Different perspectives on the Bayes factor

Jorge N. Tendeiro

Hiroshima University

09-Dec-2021

Email: tendeiro@hiroshima-u.ac.jp
Webpage: www.jorgetendeiro.com

Today

- What is the Bayes factor?

- What did I think of it?
  Tendeiro and Kiers (2019)
- vRW trying to come to the rescue.
  van Ravenzwaaij and Wagenmakers (2021)
- Nope, I don't think so.
  Tendeiro and Kiers (2021)

What is the Bayes factor?

What is the Bayes factor?

To put it lightly:

> *The Bayes factor (BF) is the Bayesian way of testing hypotheses.*

To put it accurately:

> *The BF compares the predictive ability of two competing models.*

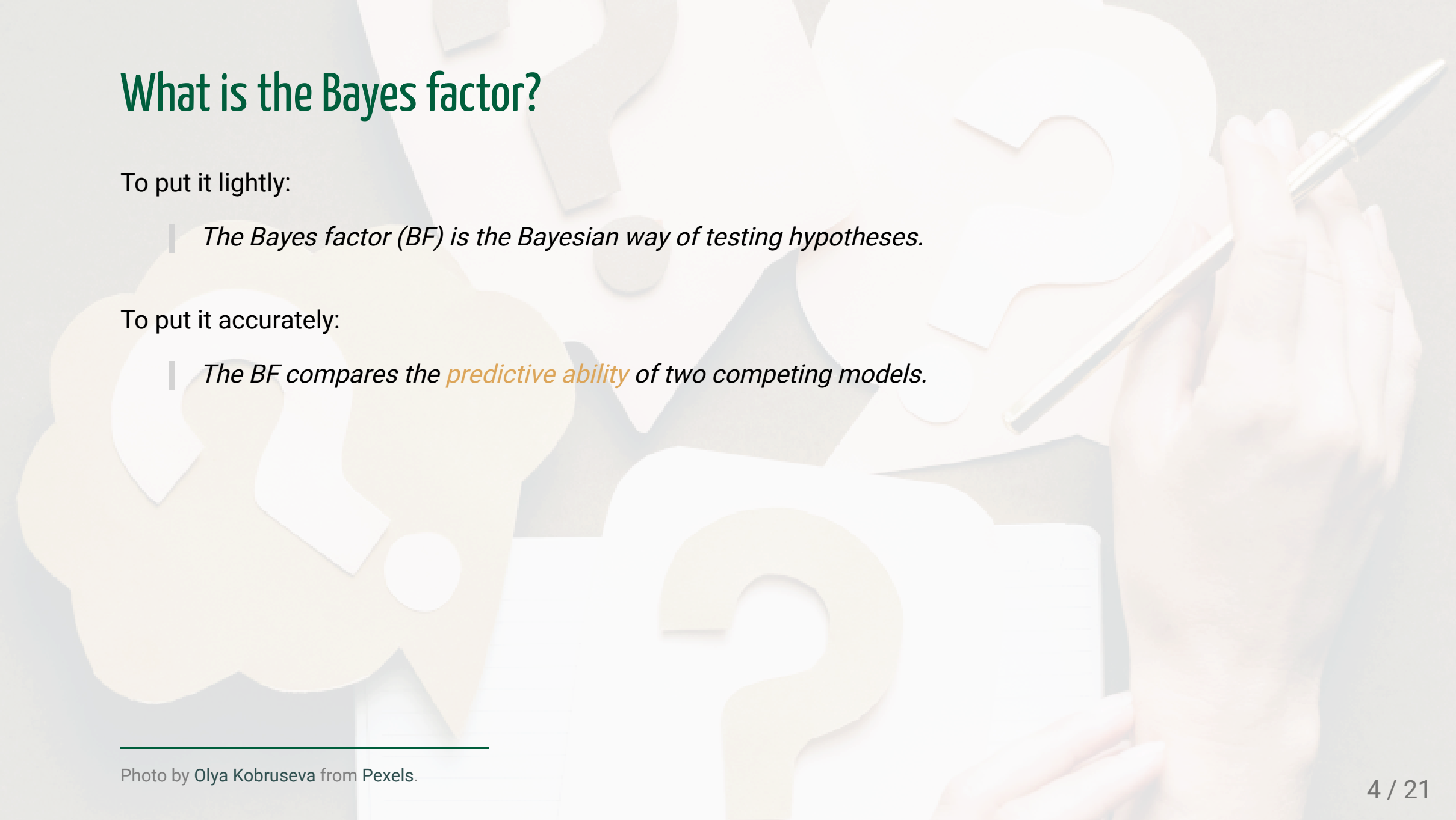


Photo by Olya Kobruseva from Pexels.

What is the Bayes factor?

From Bayes rule we can derive that

$$\underbrace{\frac{p(\mathrm{M}_0)}{p(\mathrm{M}_1)}}_{\text{prior odds}} \times \underbrace{\frac{p(D \mid \mathrm{M}_0)}{p(D \mid \mathrm{M}_1)}}_{\text{Bayes factor, } BF_{01}} = \underbrace{\frac{p(\mathrm{M}_0 \mid D)}{p(\mathrm{M}_1 \mid D)}}_{\text{posterior odds}}$$

Two possible interpretations for BF_{01}:

1. It conveys the relative probability of the observed data under either model.
   In other words, it indicates what the relative predictive ability of the two models is.

2. It indicates how an *a priori* relative model belief should be rationally updated, in light of the observed data.

GIF by Achievement Hunter from Giphy.

What did I think of it?

Tendeiro and Kiers (2019)

What did I think of it? *Tendeiro and Kiers (2019)*

Why not scrutinize the BF?

> *Any methodological approach has its advantages as well as its drawbacks and NHBT is no different.* (TK2019, p. 775)

Talk about humility in science:

> *One year before this writing, the authors felt they did not know enough about NHBT to properly understand how it works, what its merits are, and what its potential limitations are. Therefore, we carried out an extensive literature study on NHBT and related topics. This article is the result of putting together a range of discussions on NHBT.* (TK2019, p. 775)

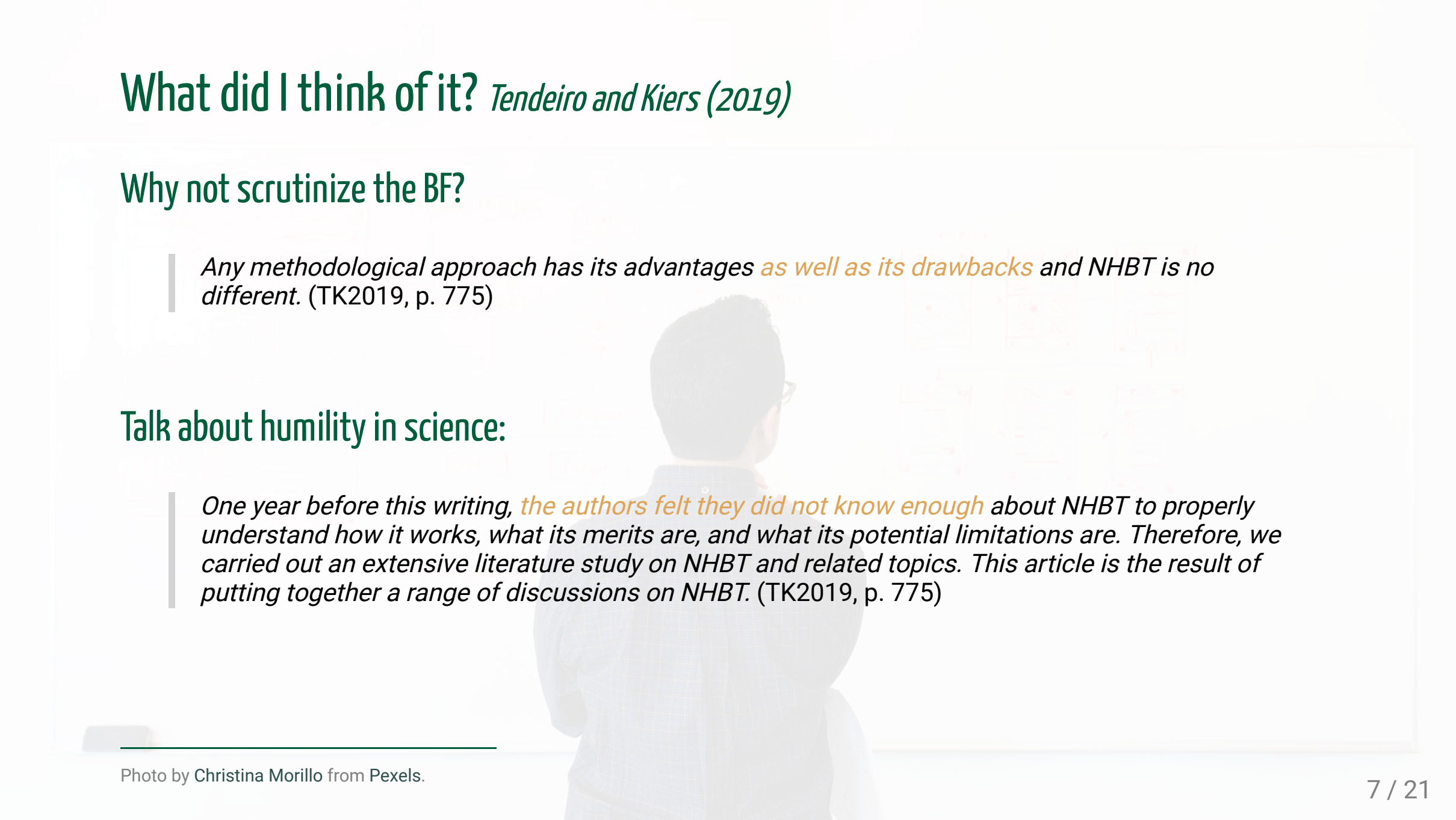


Photo by Christina Morillo from Pexels.

What did I think of it? *Tendeiro and Kiers (2019)*

The semantics of what's an "issue":

> *In this article we offer a wide overview of issues about NHBT which is currently missing in the literature.* (TK2019, Abstract)

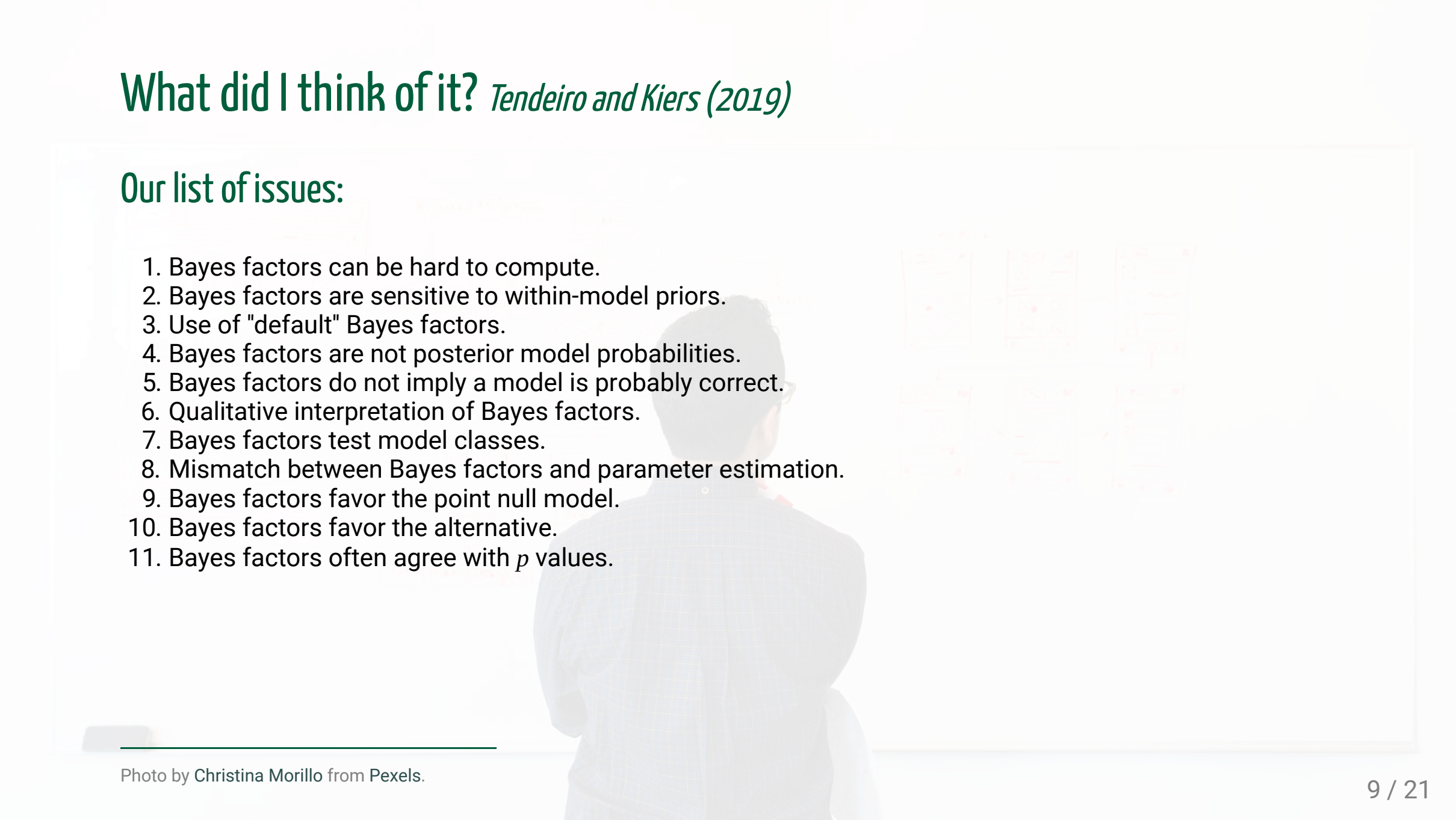
> *An "issue" can either be a limitation (according to us) or a feature that may (according to us) increase the risk of misuse or misinterpretation of a Bayes factor.* (TK2019, p. 775)

Photo by Christina Morillo from Pexels.

What did I think of it? *Tendeiro and Kiers (2019)*

Our list of issues:

1. Bayes factors can be hard to compute.
2. Bayes factors are sensitive to within-model priors.
3. Use of "default" Bayes factors.
4. Bayes factors are not posterior model probabilities.
5. Bayes factors do not imply a model is probably correct.
6. Qualitative interpretation of Bayes factors.
7. Bayes factors test model classes.
8. Mismatch between Bayes factors and parameter estimation.
9. Bayes factors favor the point null model.
10. Bayes factors favor the alternative.
11. Bayes factors often agree with p values.



vRW trying to come to the rescue

van Ravenzwaaij and Wagenmakers (2021)

vRW trying to come to the rescue *van Ravenzwaaij and Wagenmakers (2021)*

From the Abstract:

> *But although we agree with many of their thoughtful recommendations, we believe that Tendeiro and Kiers are overly pessimistic, and that several of their 'issues' with NHBT may in fact be conceived as pronounced advantages.*

> *We (...) end with a critical discussion of one of the recommendations by Tendeiro and Kiers, which is that "estimation of the full posterior distribution offers a more complete picture" than a Bayes factor hypothesis test.*

On the 11 issues:

> *In our opinion, many of the 'issues' listed by TK are a blessing rather than a curse.*
> (vRW2021, pp. 2-3)

Photo by Klaus Nielsen from Pexels.

vRW trying to come to the rescue *van Ravenzwaaij and Wagenmakers (2021)*

Their take on our 11 issues:

✗ *indicates a disadvantage;* = *indicates a neutral property;* ✔ *indicates an advantage.*

		TK	vRW
1.	Bayes factors can be hard to compute	✗	=
2.	Bayes factors are sensitive to within-model priors	✗	✔
3.	Use of 'default' Bayes factors	✗	✔
4.	Bayes factors are not posterior model probabilities	✗	✔
5.	Bayes factors do not imply a model is probably correct	✗	=
6.	Qualitative interpretation of Bayes factors	✗	=
7.	Bayes factors test model classes	✗	✔
8.	Mismatch between Bayes factors and parameter estimation	✗	=
9.	Bayes factors favor the point null model	✗	✔
10.	Bayes factors favor the alternative	✗	=
11.	Bayes factors often agree with p-values	✗	=

Photo by cottonbro from Pexels.

Nope, I don't think so

Tendeiro and Kiers (2021)

So what did we make of it?

Me and Henk thought long and hard about the reply by vRW to our paper.

> *Could it be that they are right?*

But, really, in our very *biased* opinion:

- We offered completely unassailable arguments that were not refuted by vRW.
- vRW have points of view with which we cannot concur *in the least*.

GIF by 1091 from Giphy.

Bits of what vRW had to say

E.g., concerning the sensitivity of the BF to the priors, this is both ideal:

> *The reason that Bayes factors are sensitive to within-model priors is that Bayes factors evaluate models by the predictions they make, and predictions are determined partly by the prior.* (vRW2021, p. 6),

something to be careful about:

> *(...) sensitivity analysis, that is, a comprehensive investigation of the extent to which the Bayes factor differs across alternative specification of the prior distribution.* (vRW2021, p. 13),

something that may be avoided by means of relying on defaults:

> *(...) default Bayes factors allow efficient communication of evidence that many researchers may consider less sensitive to human bias than more subjective or 'informed' alternatives.* (vRW2021, p. 13),

and it's even not that bad to close your eyes to it:

> *In our opinion, a default Bayes factor, however blindly applied, is usually preferable over no Bayes factor at all.* (vRW2021, p. 13)

Bits of what vRW had to say

Advising researchers to use BFs, even if only based on a superficial working knowledge:

> *We believe that NHBT, even if executed as a thoughtless ritual, still markedly improves on the status quo.* (vRW2021, p. 33)

About the truth of a point null hypothesis:

It most likely should be false:

> *(...) it should be stressed that the Bayes factor is based on a comparison of predictive performance that is independent of the notion of absolute or relative model truth (Wagenmakers, Grünwald, & Steyvers, 2006). The idea that we should not use models that we know to be false can easily result in an inferential impasse, because all statistical models are ultimately false.* (vRW2021, p. 26)

but at the same time it is, purportedly, a common research goal:

> *[Estimation] assumes the falsity of the null hypothesis, which is often the very target of inference.* (vRW2021, p. 34)

This is us (Tendeiro and Kiers, 2021)

So we still have something to add to the discussion:

> *Tendeiro, J. N. and Kiers, H. A. L. (2021). On the white, the black, and the many shades of gray in between: Our Reply to van Ravenzwaaij and Wagenmakers (2021). Preprint: 10.31234/osf.io/tjxvz.*

GIF from Giphy.

Some give-aways (I) *(Tendeiro and Kiers, 2021)*

- Bayes factors aren't easy to understand.
- Priors, as part of the model being compared, warrant careful attention.
- Disentangle BFs from posterior model odds.
- Relative model assessment $\neq$ truthfulness of either model.
- Conclusions derived from a Bayes factor are *always* relative, thus avoid absolute conclusions.
  So, no such thing as absolute evidence for the null!

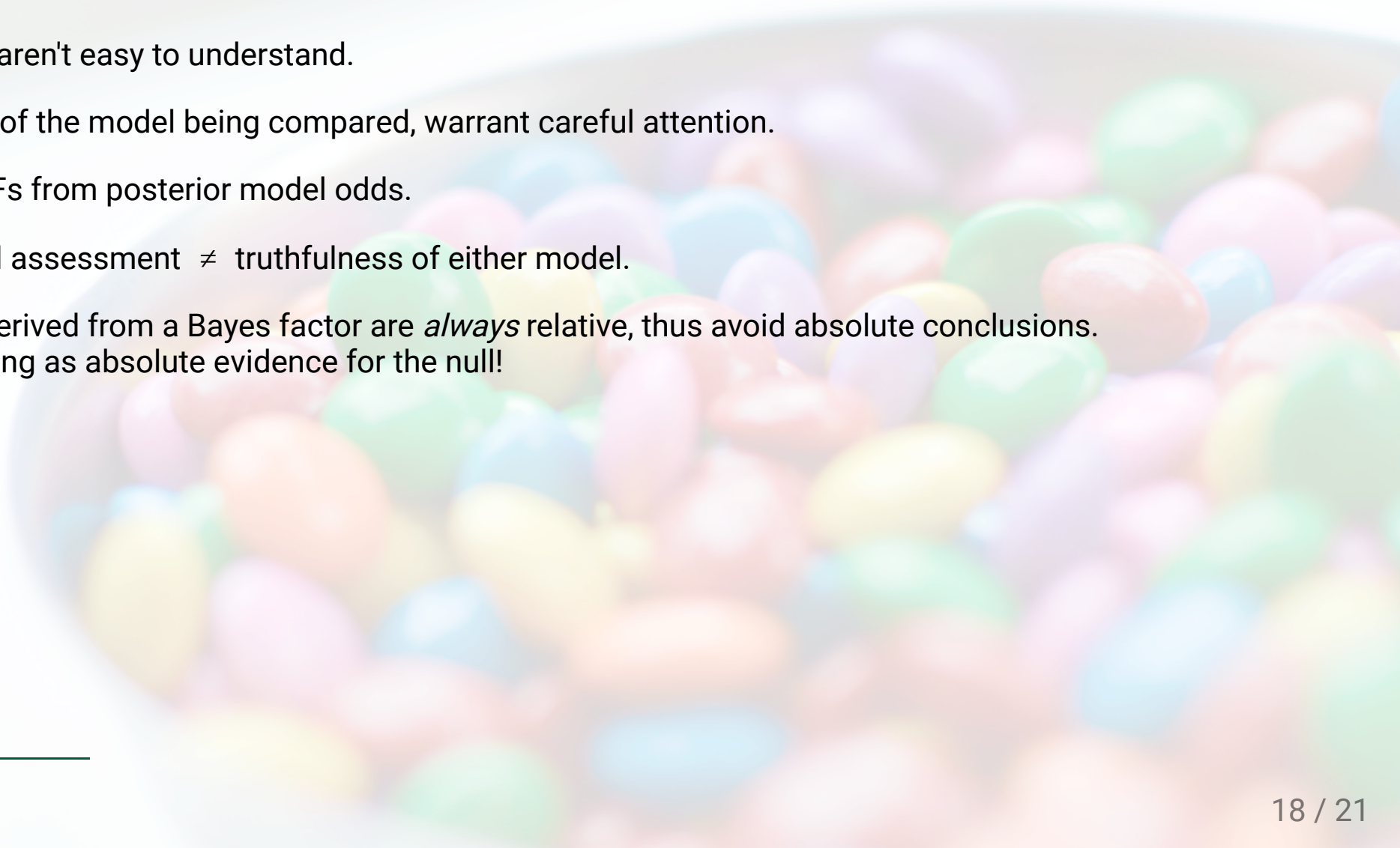


Photo by Caio from Pexels.

Some give-aways (II) *(Tendeiro and Kiers, 2021)*

- There's no such thing as the $\neq 0$ alternative hypothesis.
  There's a prior that goes with it!
- Be aware of relying on qualitative labels for BFs.
- Bayes factors can easily point at the wrong model (if you're into that kind of thing), especially for small n and ES.
- Bayes factors, in its default shape, favor the point null model.
  See our worked out example.

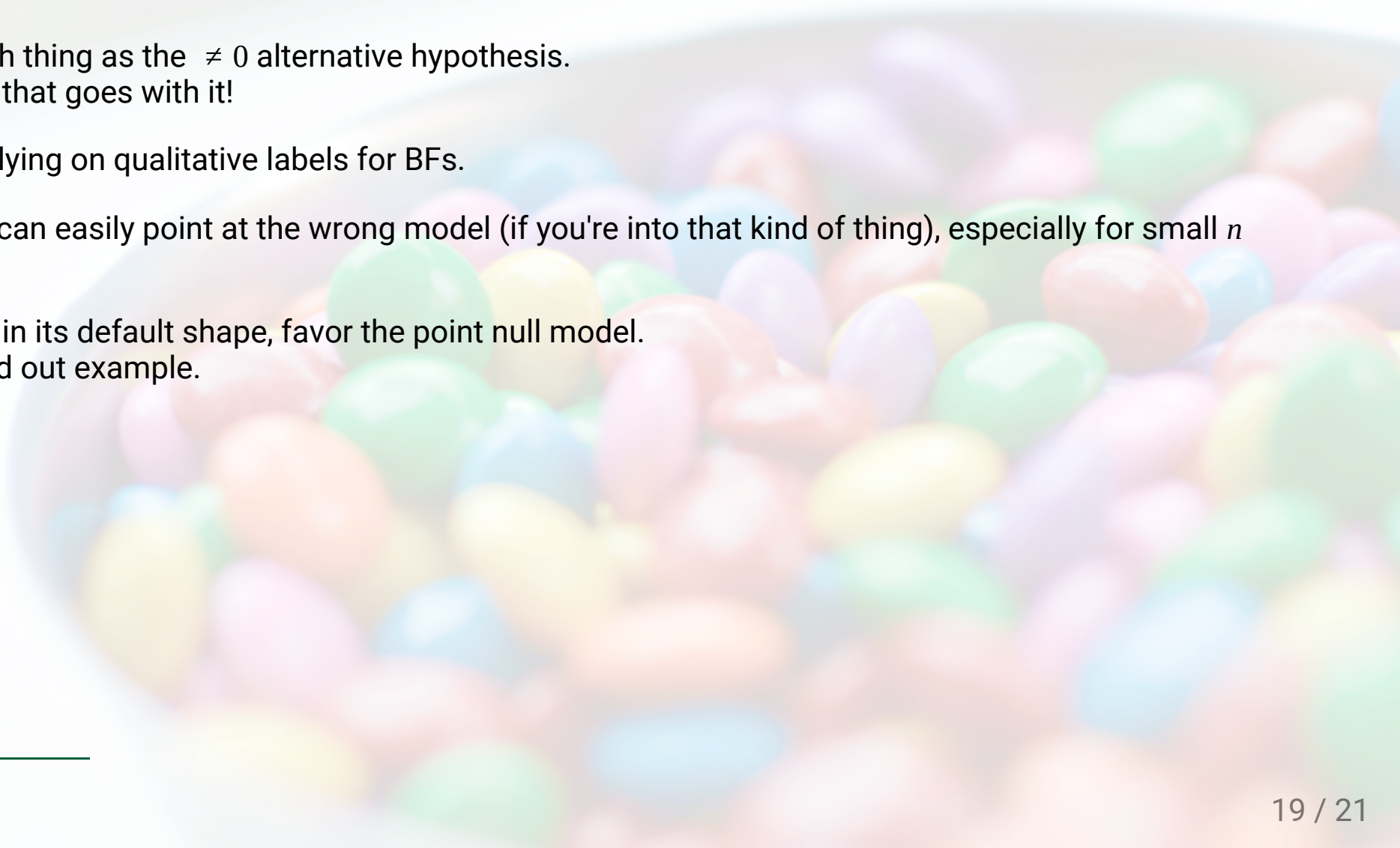


Photo by Caio from Pexels.

Some give-aways (III) *(Tendeiro and Kiers, 2021)*

- No blind endorsing of Bayes factors. Please.
- Point null hypothesis, really? At least think about it for a bit.
- Bayes factors *cannot* be used to **establish** the existence of an effect!
- We clearly reject the "first test, then who knows estimate" pseudo-inferential algorithm.
- Only reporting Bayes factor is *really* poor practice.
  Quantify effects sizes as much as you can! Estimate!

And remember...

The Bayes factor is nothing but a number, guys!

Photo by Caio from Pexels.

References

Ravenzwaaij, D. van and E. Wagenmakers (2021). *Advantages Masquerading as `Issues' in Bayesian Hypothesis Testing: A Commentary on Tendeiro and Kiers (2019)*. Preprint. PsyArXiv. DOI: 10.31234/osf.io/nf7rp.

Tendeiro, J. N. and H. A. L. Kiers (2019). "A Review of Issues about Null Hypothesis Bayesian Testing." In: *Psychological Methods* 24.6, pp. 774-795. DOI: 10.1037/met0000221.

Tendeiro, J. N. and H. A. L. Kiers (2021). *On the White, the Black, and the Many Shades of Gray in between: Our Reply to van Ravenzwaaij and Wagenmakers (2021)*. DOI: 10.31234/osf.io/tjxvz.