# Is Reflective Equilibrium too Conservative?

Andreas Freivogel

Institute of Philosophy, University of Bern



#### What is Reflective Equilibrium?

- Prominent method of justification in many subfields of philosophy ranging from ethics to logic
- ▶ Proposed by Goodman (1955) and coined "reflective equilibrium" (RE) by Rawls (1971)
- ► Key components:
- commitments ("judgments") and elements of a theory ("principles")
- ▷ starting from initial commitments, a *process* of mutual adjustments going back and forth between commitments and theory
- ▷ a *state* of equilibrium among commitments and theory, commonly characterized by coherence

## **Generating Data**

Simulations with a Python implementation of the formal RE model Ensemble: 11250 model setups with dialectical structures  $(3 \cdot 75)$ , initial commitments (25) and configurations of weights (2).

#### Distance



## The Objection of Conservativity

- ▶ longstanding line of criticism against RE(e.g., Singer (1974), Kelly and McGrath (2010)) and acknowledged by proponents of RE (e.g., Scanlon (2003))
- Coherence considerations in RE do not provide enough incentive to revise the initially held views substantially.
- Consistency and "hanging together" can be established easily by streamlining the starting point.
- Conservativity is a pressing issue in view of epistemically deficient inputs (biases, prejudices): "garbage in – garbage out".

## A Formal Model of RE

► Based on elaborate accounts of RE, Beisbart et al. (2021) provide a formal model in a propositional framework.





#### **Inconsistency Preservation**



Figure: The model performs worse/better in (a)/(b) than the simplistic adaptation baseline.



Figure: Components of the formal model of RE.

- Overall achievement of an epistemic state is defined as
  - $Z(C, T \mid C_0) = \alpha_A \cdot A(C, T) + \alpha_S \cdot S(T) + \alpha_F \cdot F(C \mid C_0)$
- $\blacktriangleright$  RE states are global optima  $\Box$  according to Z that satisfy additional requirements.
- $\triangleright$  Z is also utilized to give explicit rules for equilibration processes that yield fixed points .
- Trade-offs between the desiderata of account, systematicity and faithfulness in Z are determined by a configuration of weights

# **Axiomatic Simplicity**



Figure: The RE model performs similar/better in (a)/(b) than the streamlining baseline.

#### **Conclusion: RE is not too Conservative**

► The formal model performs significantly better than conservative baselines For specific configurations of weights.

#### $(\alpha_A, \alpha_S, \alpha_F).$

## **Conservative Baselines**

- Streamlining : Find a minimal axiomatic base of the initial commitments. Fit the commitments perfectly to the consequences of the axiomatic base.
- Simplistic adaptation : Keep the initial commitments if they are consistent. Otherwise, give up a random commitment.
- ► Compare model performance ( , ) baselines with respect to operationalizable features that would reveal conservative behaviour: distance, inconsistency preservation, and axiomatic simplicity

- ► In general, well-performing configurations put a a lot of weight to the desiderata of account  $(\alpha_A)$  and systematicity  $(\alpha_S)$  that operationalize theoretical virtues in the model.
- Including and stressing the importance of theoretical virtues in RE may help to overcome the objection of conservativity.

## References

Claus Beisbart, Gregor Betz, and Georg Brun. Making reflective equilibrium precise. a formal model. Ergo, 8(15):441-472, 2021. Nelson Goodman. Fact, Fiction, and Forecast. Cambridge, Massachusetts: Harvard University Press, 1955. Thomas Kelly and Sarah McGrath. Is reflective equilibrium enough? Philosophical Perspectives, 24(1):325-359, 2010. John Rawls. A Theory of Justice. Cambridge, Massachusetts: The Belknap Press of Harvard University Press, 1971. Thomas M. Scanlon. Rawls on justification. In Samuel R. Freeman, editor, The Cambridge Companion to Rawls, pages 139–167. Cambridge University Press, 2003.

Peter Singer. Sidgwick and reflective equilibrium. The Monist, 58(3):490-517, 1974

How far does reflective equilibrium take us? Investigating the power of a philosophical method (SNF grant no. 182854, DFG grant no. 412679086)

andreas.freivogel@philo.unibe.ch