What Does Rational Analysis Tell Us about Rationality?

Anderson 1990, Chapters 1 and 6

Kevin Dorst
kevindorst@pitt.edu

Rationality Seminar
February 1, 2021

I. The Question

Potted history: Heuristics and biases research program was dominant in 1970s and 80s. It evolved into "dual-process" theories (System 1 vs. System 2) and behavioral economics.

Kahneman et al. 1982
Kahneman 2011
Thaler 2015

But starting with Anderson's book, a "rational analysis" or "Bayesian" approach to cognition began gaining momentum. Now it's a flourishing (if controversial) research program.

Oaksford and Chater 1998, 2007
Griffiths et al. 2008; Tenenbaum et al. 2011; Lewis et al. 2014

Question: How do we square these two research programs?

Do they compete?
If not, why not?
If so, which is right?

II. Levels of analysis

Could a neuroscientist understand a microprocessor?

Jonas and Kording 2017

Marr on understanding information-processing systems:

1) Computational level: What is the system doing, and why?
2) Algorithmic level: How does it do it? (What program is it running?)
3) Implementational level: how is that program implemented in the brain?

Vision: infer 3D scene from 2D array.
Identify contour lines and shade gradients; use parallax.
???

(Anderson changes"computational" to "rational"; "implementational" to "biological"; and splits "algorithmic" into "algorithm" and "implementation" levels. Don't worry too much about this.)

Much of psychology focuses on (2)—concern with *mechanism*.

Think of "heuristics" like representativeness, availability, or anchoring.

Why need? Feathers and aerodynamics (Marr).

Anderson: likewise human memory and information-retrieval systems!

Anderson: core is **adaptationist assumption**: the system is optimized to solve a particular problem.

Both Marr and Anderson bemoan "grab bag" mechanistic approach.

Anderson's method:

1) Specify goal of cognitive system.
2) Develop formal model of environment in which it adapted.
3) Make minimal assumptions assumptions about computational costs/limitations.
4) Derive optimal behavior, given these.
5) Compare predictions to empirical data.
6) Iterate.

Example: memory. Goal is to recall relevant information. Cues + past usages yield need probabilities. Constant cost of search. Search through linearly (in order of need-probability) until either find, or expected value is lower than cost of computation.
Explains: power-law of learning; spacing effects; priming effects; fan- and expertise-effects, etc...

This has been used—with a fair amount of empirical success—to help explain memory, categorization, causal inference, and problem-solving (Anderson); but also in areas that H&B had purportedly shown systematic biases, e.g. prediction, confirmation, and influences of priors (base rate neglect or conservatism).

Tenenbaum and Griffiths 2006
Oaksford and Chater 1994, 2003; Austerweil and Griffiths 2011
Dasgupta et al. 2019

The degree of this success is controversial, but lets suppose the standard success narrative is right.

Bowers and Davis 2012; Griffiths et al. 2012

Big Question: What would that mean for human rationality?

How do we square the success of heuristics-and-biases tradition with success of rational-analysis tradition?

First: why even give rational analysis a seat at the table?

II. Arguments for rational analysis

Methodological and epistemological.

1) Empirical success.

Let's grant them that.

2) Problem of identifiability (underdetermination)

The empirical data we're able to gather about what algorithms are being computed and how that happens is very limited.

Anderson focuses on his implementational (vs. algorithmic level); not totally sure why it matters.

By the plenitude of algorithms, this leads to *massive* underdetermination.

Without a guiding (computational-level) principle in selection, this is like shooting in the dark.

Upshot: we shouldn't lend much credence to the mechanisms this process (like H&B?) generates.

Advantages of rational analysis:

Optimal given environment and goals.

- Environment and goals much more accessible.
- Optimality puts heavy constraints on choice of mechanism.
- If true, gives better explanation. (*Why?*)

Easier to see outside the head than inside it!

Q: Thoughts?

Is there an *alternative* computational-level explanation, besides optimality?

3) Evolutionary argument

Evolution is a *local optimizer*. Hill-climbing metaphor.

Note: *not* a satisficer. Intra-species competition.

Evolutionary explanations are better the less arbitrary the "fitness landscape" is. (And presumably: the more we know about it.)

Shape of feathers vs. color (moths).

Leads to good explanation for many important biological features (bird flight; human vision; arterial structure).

Cognition has relatively robust fitness landscape. What to remember, how to categorize, or make causal inferences, or predict, or test hypotheses. In all these cases, the fitness landscape (how to do it well vs. poorly) is fairly robust.

Depends? Memory or categorization vs. causal inference. Might be robust in evolutionary past but not today.

Q: Thoughts?

Also depends on *size* of space and *time* to hill-climb. Intractability of central cognition? Language-based distortions of experience-sampling?

III. Friends or Foes?

Is rational analysis *in competition* with heuristics-and-biases? If so, what to make of that fact?

Anderson gives several different reactions. By-and-large, resulting literature has concluded that there's a tension.

Friends?

1) Different senses of 'rational'?

1.i) Normative vs. *ecological* rationality?

Epistemic ≈ follow rules of epistemology/logic/decision theory.

Nod to Cohen: the *right* rules, not just those in the textbooks.

Ecological ≈ in evolutionary environment, take best strategy for producing offspring.

a) Offspring goal ≠ epistemic goal?
b) Evolutionary environment differs systematically?

Worry: For core cognition, how much daylight is there between these two?

- (Claimed) robustness of fitness landscape contradicts (b).
- Plausibly, best way to produce offspring is to get accurate beliefs, take good means to your ends, and *instill the right ends*.
- In particular: was representativeness (ignoring base rates) really a good way to make predictions in our evolutionary past?

Exceptions? E.g. experience-sampling and language?

1.ii) Ideal vs. "bounded" rationality?

Perhaps key is that rational analysis builds in computational costs and limitations, and the adaptive costs of violating normative rules are often low?

E.g. primacy effect

Worry: H&B didn't want to show deviations from *ideal* rationality—we all already knew that! (We don't play chess perfectly.) And they want to say the errors the heuristics generates are *serious*.

Remember the grant-writing...

2) Different cognitive systems under different selection pressures?

Vision and memory are highly constrained problem spaces. H&B cases of apparent rationality, like prediction, not so.

Certainly some truth to this!

Worries:

- Even so, surely something as basic as base-rate neglect could easily be selected out!
- In practice, rational analysis and H&B often compete for explanations of same phenomena.

Note: "conservatism" results as well.

Foes?

1) H&B have wrong normative standards

Expected value across *range* of normal cases.

E.g. causes similar to effects
Note: *any* approximation will generate "systematic" deviations; just find the circumstances where approx ≠ target.

Undersell uncertainty in subject's situation. (Get in their head!)

Bad assumptions about goals of the system?

Memory for quickly retrieving from lifetime, not for storing 20 digits for 5 minutes.

2) Rational analysis must be wrong!

E.g. because math too hard.

Anderson: no assumptions about computations the brain is doing. *Others:* people's brains do intractable calculations *all the time*.

Is that right?

References

Austerweil, Joseph L and Griffiths, Thomas L, 2011. 'Seeking Confirmation Is Rational for Deterministic Hypotheses'. *Cognitive Science*, 35:499–526.

Bowers, Jeffrey S. and Davis, Colin J., 2012. 'Bayesian just-so stories in psychology and neuroscience'. *Psychological Bulletin*, 138(3):389–414.

Dasgupta, Ishita, Schulz, Eric, Tenenbaum, Joshua B., and Gershman, Samuel J., 2019. 'A Theory of Learning to Infer'.

Griffiths, Thomas L., Chater, Nick, Norris, Dennis, and Pouget, Alexandre, 2012. 'How the bayesians got their beliefs (and what those beliefs actually are): Comment on bowers and davis (2012)'. *Psychological Bulletin*, 138(3):415–422.

Griffiths, Thomas L., Kemp, Charles, and Tenenbaum, Joshua B., 2008. 'Bayesian Models of Cognition'. In R. Sun, ed., *The Cambridge Handbook of Computational Psychology*, 59–100. Cambridge University Press.

Jonas, Eric and Kording, Konrad Paul, 2017. 'Could a Neuroscientist Understand a Microprocessor?' *PLoS Computational Biology*, 13(1):1–24.

Kahneman, Daniel, 2011. *Thinking Fast and Slow*. Farrar, Straus, and Giroux.

Kahneman, Daniel, Slovic, Paul, and Tversky, Amos, eds., 1982. *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge University Press.

Lewis, Richard L., Howes, Andrew, and Singh, Satinder, 2014. 'Computational rationality: Linking mechanism and behavior through bounded utility maximization'. *Topics in Cognitive Science*, 6(2):279–311.

Oaksford, Mike and Chater, Nick, 1994. 'A Rational Analysis of the Selection Task as Optimal Data Selection'. *Psychological Review*, 101(4):608–631.

———, 1998. *Rational models of cognition*. Oxford University Press Oxford.

———, 2003. 'Optimal data selection : Revision , review , and reevaluation'. *Psychological Bulletin & Review*, 10(2):289–318.

———, 2007. *Bayesian rationality: The probabilistic approach to human reasoning*. Oxford University Press.

Tenenbaum, Joshua B and Griffiths, Thomas L, 2006. 'Optimal Predictions in Everyday Cognition'. *Psychological Science*, 17(9):767–773.

Tenenbaum, Joshua B, Kemp, Charles, Griffiths, Thomas L, and Goodman, Noah D, 2011. 'How to grow a mind: Statistics, structure, and abstraction'. *science*, 331(6022):1279–1285.

Thaler, Richard H., 2015. *Misbehaving: The Making of Behavioural Economics*. Penguin.