Behavioral and Brain Sciences

cambridge.org/bbs

# Conviction Narrative Theory: A theory of choice under radical uncertainty

Samuel G. B. Johnson[a,b,c,d], Avri Bilovich[b] and David Tuckett[b,e]

[a]Department of Psychology, University of Warwick, Coventry CV4 7AL, UK; [b]Centre for the Study of Decision-Making Uncertainty, University College London, London W1CE 6BT, UK; [c]University of Bath School of Management, Bath BA2 7AY, UK; [d]Department of Psychology, University of Waterloo, Waterloo, ON N2L 3G1, Canada and [e]Blavatnik School of Government, University of Oxford, Oxford OX2 6GG, UK.
sgbjohnson@gmail.com
a.bilovich@ucl.ac.uk
d.tuckett@ucl.ac.uk

**Target Article**

**Cite this article:** Johnson SGB, Bilovich A, Tuckett D. (2023) Conviction Narrative Theory: A theory of choice under radical uncertainty. *Behavioral and Brain Sciences* **46**, e82: 1–74. doi:10.1017/S0140525X22001157

Target Article Accepted: 20 May 2022
Target Article Manuscript Online: 30 May 2022
Commentaries Accepted: 13 September 2022

**Keywords:**
behavioral economics; causal thinking; cognitive science; decision-making; emotion; explanations; imagination; narratives; reasoning; social learning; uncertainty

**What is Open Peer Commentary?** What follows on these pages is known as a Treatment, in which a significant and controversial Target Article is published along with Commentaries (p. 26) and an Authors' Response (p. 65). See bbsonline.org for more information.

**Abstract**

Conviction Narrative Theory (CNT) is a theory of choice under *radical uncertainty* – situations where outcomes cannot be enumerated and probabilities cannot be assigned. Whereas most theories of choice assume that people rely on (potentially biased) probabilistic judgments, such theories cannot account for adaptive decision-making when probabilities cannot be assigned. CNT proposes that people use *narratives* – structured representations of causal, temporal, analogical, and valence relationships – rather than probabilities, as the currency of thought that unifies our sense-making and decision-making faculties. According to CNT, narratives arise from the interplay between individual cognition and the social environment, with reasoners adopting a narrative that feels "right" to explain the available data; using that narrative to imagine plausible futures; and affectively evaluating those imagined futures to make a choice. Evidence from many areas of the cognitive, behavioral, and social sciences supports this basic model, including lab experiments, interview studies, and econometric analyses. We identify 12 propositions to explain how the mental representations (narratives) interact with four inter-related processes (explanation, simulation, affective evaluation, and communication), examining the theoretical and empirical basis for each. We conclude by discussing how CNT can provide a common vocabulary for researchers studying everyday choices across areas of the decision sciences.

> Before the wheel was invented… no one could talk about the probability of the invention of the wheel, and afterwards there was no uncertainty to discuss…. To identify a probability of inventing the wheel is to invent the wheel.
>
> John Kay and Mervyn King, *Radical Uncertainty* (2020)

## 1. Everyday decisions

A government amidst a public health lockdown debates exit strategy; a couple debates divorce. A university graduate considers her career options; the CEO of a toaster company considers expanding into blenders. A widow, awoken by a strange sound, contemplates whether to investigate its source; a burglar, outside, contemplates whether he is making a grave mistake.

We make such decisions, grand and petite, every day. This is remarkable because many everyday choices require us to solve six challenges – each daunting, together herculean:

- **Radical uncertainty.** *Our knowledge about the future often eludes quantification*. (Experts give conflicting advice to the government; the bickering couple cannot know whether their past signals their future.)
- **Fuzzy evaluation.** *The criteria for evaluating the future are ambiguous and multidimensional.* (The couple must consider their feelings, children, finances; careers bring different forms of satisfaction.)
- **Commitment.** *Decisions and outcomes are often separated in time, so we must manage our course of action as the situation evolves.* (People must sustain career training and organize their plans for years on end.)
- **Sense-making.** *The right decision about the future depends on grasping the present.* (The government considers which epidemiological models are most plausible; the widow makes her best guess about what caused the noise.)
- **Imagination.** *Since the future does not yet exist, we must imagine it to evaluate its desirability*. (Decisions about love, appliances, intruders, and viruses require future forecasts.)

- **Social embeddedness.** *The decision depends both on our beliefs and values, and those of others.* (The government persuades the public to implement its policies; beliefs about marriage are shaped by our culture and media diet.)

These challenges are ubiquitous, yet their solutions elude dominant theories of decision-making.

This paper presents *Conviction Narrative Theory* (CNT) – an account of choice under radical uncertainty. According to CNT, *narratives* – mental representations that summarize relevant causal, temporal, analogical, and valence information – are the psychological substrate underlying such decisions. Narratives support and link four processes – *explanation* (structuring evidence to understand the past and present, yielding emotional satisfaction), *simulation* (generating *imagined futures* by running the narrative forward), *affective evaluation* (appraising the desirability of imagined futures and managing commitment toward a course of action over time), and *communication* (transmitting decision-relevant knowledge across social networks to justify, persuade, and coordinate action). Narratives are why the above-mentioned properties so often co-occur: In contexts marked by radical uncertainty and fuzzy evaluation, we use narratives to make sense of the past, imagine the future, commit to action, and share these judgments and choices with others.

Narratives bubble beneath every example above. Governments debate whether a virus is more like flu or plague; these narratives yield very different explanations of the situation, hence predictions about the future, hence emotional reactions to particular options. The couple can interpret their fights as signaling differences in fundamental values or resulting from temporary stresses; either narrative can explain the fights, portending either a dark or rosy future. The toaster CEO might consider her company ossified, complacent, or innovative; these narratives have different implications about the risks and benefits of new ventures, motivating different decisions. In each case, the decision-maker's first task is to understand the current situation, which informs how they imagine a particular choice would go, which is deemed desirable or undesirable based on how the decision-maker would feel in that imagined future.

Narratives pervade decision-making. This article explains why and how.

Samuel G. B. Johnson, Assistant Professor of Psychology at the University of Waterloo and Visiting Fellow at the University of Bath School of Management, has published over 50 journal articles and conference papers on inference, decision-making, moral judgment, and consumer behavior. He received the Glushko Dissertation Prize in 2017 from the Cognitive Science Society for his research on explanatory reasoning.

Avri Bilovich, Honorary Research Fellow at the Centre for the Study of Decision-Making Uncertainty at UCL.

David Tuckett, Emeritus Professor at University College London (UCL) and Senior Research fellow at the Blavatnik School of Government, University of Oxford, is the author of publications in the fields of economics, economic history, decision-making, medical sociology, and psychoanalysis, including *Minding the Markets: An Emotional Finance View of Financial Instability*, and has received the Sigourney Award for his contribution to psychoanalysis in both 2008 and 2019.

## 2. The logic of decision

### 2.1. Two problems

Any theory of decision-making must account for how beliefs and values yield action. We divide this question into two problems – *mediation* and *combination*.

#### 2.1.1. The mediation problem

Since data must be interpreted to be useful, decision-making requires a mental representation – a *currency of thought* – mediating between the external world and our actions (dashed lines in Fig. 1). When we face a decision, we form beliefs – based on prior knowledge and new data – to characterize what will likely happen given potential actions. Those beliefs must be represented in a format that can be combined with our values to guide action (Baumeister & Masicampo, 2010). Put differently, external *data* or raw facts do not become actionable *information* until interpreted in conjunction with our broader knowledge (Tuckett, Holmes, Pearson, & Chaplin, 2020). The burglar must consider, were he to burgle, the likely outcomes (beliefs) and their desirability (values). Some mental representation must simultaneously be the *output* of the reasoning process that judges what will happen and an *input* to the decision-making process that combines those beliefs with each outcome's desirability.

In classical decision theory, the currency of thought is *probability* – continuous values that quantify risk. The burglar's decision depends on his perceived chance he will be caught ($C$) or not caught ($NC$). But this assessment depends potentially on many things – the police presence, burglar's skill, odds the inhabitants are home, etc. These data must be aggregated through Bayesian inference (Section 4.1), combining prior knowledge with new evidence.

The burglar weighs the evidence, assigning 0.2 probability to $C$ and 0.8 to $NC$. These probabilities summarize all relevant data about the *external* world in a format used *internally* to combine these beliefs about outcomes with values about their desirability. Probabilities solve the mediation problem because a single representation can be the output of belief-formation and the input to decision-making.

#### 2.1.2. The combination problem

Decision-making requires a process – a *driver of action* – that combines beliefs and values to yield action. The burglar must not only assess the likelihood of being caught or not, but how bad or good that would be. If the mediation problem has been solved, we have a suitable representation of likelihood to combine with value judgments. Yet a further principle must govern this combination.

In classical decision theory, the driver of action is *utility-maximization*: Disparate sources of value are aggregated into an outcome's *utility*, multiplied by each outcome's probability to yield an option's *expected utility*. The decision rule is simply to maximize this quantity. The burglar would consider the sources of (dis)utility associated with being caught ($C$) – social stigma, financial costs, prison – and with not being caught ($NC$) – newfound wealth, perhaps guilt. The utility of $C$ and $NC$ might be −8 and +3, respectively. Then, the expected utility is each state's utility, weighted by its probability:

$$U(C) \times P(C) + U(NC) \times P(NC) = (0.2)(-8) + (0.8)(3) = +0.8$$

Crime is expected to pay, so the rational burglar would attempt the burglary. Although expected utility maximization is not the only justifiable decision rule, philosophers and economists have

https://doi.org/10.1017/S0140525X22001157 Published online by Cambridge University Press

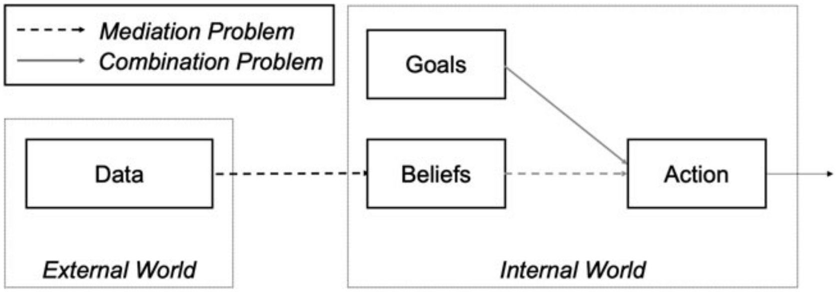


**Figure 1.** The logic of decision.
Decisions reflect both data picked up from the external world – including the social environment – and internally derived goals. The mediation problem (dashed lines) reflects the need for an internal representation – a currency of thought – that can mediate between data from the external world and actions decided internally. The combination problem (gray lines) reflects the need for a process – a driver of action – that can combine beliefs and goals to yield actions. In classical decision theory, the currency of thought is probability and the driver of action is expected utility maximization. In CNT, the currency of thought is narratives, and the driver of action is affective evaluation.

marshalled powerful arguments for its rationality (Savage, 1954; Von Neumann & Morgenstern, 1944).

### 2.2. Two puzzles

The reader may already feel uneasy about these admittedly ingenious solutions to the mediation and combination problems: *Where do these numbers come from?* Varieties of this puzzle afflict both probabilities and utilities.

#### 2.2.1. Radical uncertainty

*Radical uncertainty* characterizes situations when probabilities are unknowable (Kay & King, 2020; Knight, 1921; Volz & Gigerenzer, 2012), because we do not know the data-generating model or cannot list all possible outcomes. Debates over pandemic policy are riddled with uncertainty about the infection itself (contagiousness, lethality) and policy responses (efficacy, unintended consequences). We don't know the right model for any of these – without a model, how do we calculate probabilities? Moreover, we cannot enumerate the potential implications of each policy choice – without a list of outcomes, how do we assign probabilities to them? Similar problems afflict our other examples earlier – try assigning probabilities to the prospects of the bickering couple, the toaster CEO, or, indeed, the burglar, and it becomes clear that radical uncertainty haunts many everyday decisions.

Radical uncertainty has many sources. Some derive from *aleatory uncertainty* from the world itself (Kay & King, 2020):

- **Non-stationary distributions.** Stationary processes have constant probability distributions over time, learnable over repeated observations. Many real-world processes are *non-stationary*. Each time a pathogen mutates, its previously observed properties – the severity of disease in population sub-groups, its responsiveness to treatments, and prevention by vaccines – change in unknowable ways. The question "What is the probability of dying of a mutating virus if I contract it in 6 months?" has no answer.
- **Agency.** Human behavior is often unpredictable. This is especially obvious for pivotal historical events – the assassination of Caesar, Putin's invasion of Ukraine – but smaller forms of agency-driven uncertainty render foggy whole swaths of the future. Technological innovation depends on the insights and happenstance of individuals (Beckert, 2016; Knight, 1921; Ridley, 2020), yet produce profound discontinuities. Behavior emerging from interactions among collectives adds further uncertainty, as illustrated by waves of virality in social media. The sweeping effects of government policies often depend on the preferences of one person or unpredictable interactions among a group. The COVID-19 pandemic would have had a far different shape were it not for many unpredictable choices – the rapid development of vaccines by scientists, the often-haphazard decisions of politicians.

Radical uncertainty also results from the *epistemic* limitations of our finite minds:

- **Information limits.** Often, we lack information to fully understand a situation. In the early days of a pandemic, we know little about how a pathogen is transmitted or who is afflicted. At other times, we have *more* information than we can process: An endless parade of potentially relevant data resides in our environment and the deepest trenches of memory. There is often an abundance of relevant information, if only we knew where to find it. But life is not an exam problem – information is not branded with "relevant" or "irrelevant."
- **Specification limits.** When we do not know the data-generating model, we often cannot rationally assign precise probabilities (Goodman, 1955). This means that Bayesian inference – combining precisely expressed prior beliefs with quantitative assessments of how well the data fit each hypothesis – is often mathematically ill-posed. Much thought is instead more qualitative (Fisher & Keil, 2018; Forbus, 1984); while this can create bias, it is often unclear even normatively how to assign precise values. To generate probabilities, data must be interpreted; interpretation requires a model; and our models, for all but the simplest situations, are incomplete.
- **Generation limits.** Most realistic problems are open-ended, requiring us to generate our own hypotheses. There are endless reasons why a new cluster of virus cases can arise – a resident returned from abroad, a tourist brought the virus, a super-spreader event happened, a new variant has arisen. Even if we can *test* individual explanations, we will never be able to *list* all possible explanations. Our imaginations are limited and so we cling to small numbers of especially plausible hypotheses – raising the question of where they come from.
- **Capacity limits.** Our minds have limited attention, working memory, and inference capacity (Miller, 1956; Murphy & Ross, 1994). Bayesian calculations rapidly reach absurd calculational complexity. For each calculation of a posterior, we must separately calculate the prior and likelihood and combine them, and an inference may require posteriors for many plausible hypotheses. This is bad enough, but often our inferences are chained (Steiger & Gettys, 1972). In the case of a pandemic, we cannot generate reliable predictions of how death numbers will respond to policy interventions because the responses both of individuals (e.g., distancing behavior) and the virus (e.g., mutations) are uncertain and intertwined in feedback loops.

Probabilities, by definition, are inappropriate under radical uncertainty. Although uncertainty has long been a thorn in the side of

https://doi.org/10.1017/S0140525X22001157 Published online by Cambridge University Press

economics (Camerer & Weber, 1992; Ellsberg, 1961; Knight, 1921), almost all economic models assume that outcomes can be enumerated and assigned probabilities. Even behavioral models typically replace optimal with biased probabilistic processing (Tversky & Kahneman, 1979). This can work when the underlying model really is known, as in gambling. But real-world decision-making often resembles poker more than roulette – probabilities only get you so far.

#### 2.2.2. Fuzzy evaluation

*Fuzzy evaluation* characterizes situations in which utilities cannot be evaluated. Reasons include:

- **Incommensurable attributes.** We rarely evaluate choice objects along a single dimension, but must somehow combine multiple dimensions into an overall summary judgment. Writing an academic article mingles the joy of intellectual work and the pride of completion with the frustration of slow progress and the angst of possible rejection. Filing for divorce merges the pain of leaving behind shared history with the prospect of turning over a new leaf. These potential options are difficult to evaluate because these attributes are along almost totally unrelated dimensions that resist placement onto a common scale (Walasek & Brown, 2021).
- **Incomparable outcomes.** When we compare objects along a single dimension, we can often simply rank them pairwise along that dimension. But so often, one object excels on one dimension while another object excels on another. When the attributes are incommensurable, trading off attributes across choice objects is necessary to make a choice, yet it is often unclear how to do so rationally (Walasek & Brown, 2021). For example, consider choosing between careers as a clarinetist or lawyer (Raz, 1986). Neither career is clearly better, nor are they equally good – they are good in *different* ways: One involves more self-expression, the other more opportunity to improve the world. The relative desirability of these attributes eludes quantification. Imagine increasing the clarinetist's salary by 1%. Although clearly better than the original clarinetist job, it is still not clearly better than the lawyer job, violating transitivity (Sinnott-Armstrong, 1985). Gaining further information is unlikely to help here, where there are good arguments for and against each choice – a recipe for ambivalence (Smelser, 1998).
- **Non-stationary values.** Our values may be unstable over time, yet we often make decisions for our future selves. Innovators face the challenge that consumers may not know what they like until they actually experience it – as in Henry Ford's (apocryphal) remark that if he had asked customers what they wanted, they would have said "faster horses." We decide whether to have a first child before the experience of parenthood radically alters our priorities (Paul, 2014). Just as beliefs are uncertain when probability distributions are non-stationary, so are values uncertain when they change unpredictably. Moreover, even if one could accurately predict one's future values, how can *current* decisions be governed by *future* preferences?

Just as neoclassical and behavioral models differ in their approach to uncertainty mainly in assuming optimal versus biased probabilistic processing, their approach to preferences differs mainly in adding additional sources of utility (e.g., social utility; Fehr & Schmidt, 1999) or biases (e.g., reference-dependent preferences; Tversky & Kahneman, 1979). Such approaches are poorly suited to many everyday decisions where utilities are non-calculable and fuzzy evaluation reigns.

## 3. Conviction Narrative Theory

Conviction Narrative Theory (CNT) characterizes the social and informational context in which decision-making occurs and the cognitive and affective processes governing it. CNT provides alternative solutions to the mediation and combination problems that eschew probabilities and utilities.

Under radical uncertainty and fuzzy evaluation, decision-making requires us to extract relevant information by explaining the past, use that information to predict the future, and evaluate possible futures. CNT posits *narratives* as the key mental representation underpinning these processes: A narrative is selected that best explains the data, which is then used to imagine possible futures given potential choices, with emotional reactions to those imagined futures motivating choices – producing conviction to take sustained action (Tables 1 and 2). (For precursors, see Chong & Tuckett, 2015; Tuckett, in press; Tuckett et al., 2020; Tuckett & Nikolic, 2017.)

***Context.*** Although not every decision is taken under radical uncertainty and fuzzy evaluation – probabilities and utilities are well-suited for studying gambles typical in risky-choice experiments – these properties are common in everyday decisions that do not wear numbers on their sleeves. Despite drawing on fewer resources by avoiding probabilities and utilities, CNT draws on more resources in another sense – decisions are typically socially embedded, with beliefs and values influenced by others and subject to cultural evolution (Section 9). This often permits reasonable decision-making in the absence of probabilities and utilities.

***Representations.*** CNT posits *narratives* as structured, higher-order mental representations summarizing causal, temporal, analogical, and valence structure in a decision domain (Section 5). For example, the widow hearing the noise has different causal theories of why sounds occur at different times; draws analogies between the present case and similar situations; and keeps track of the nefarious or innocent intentions implied. This knowledge might be represented in "burglary" versus "noisy cat" narratives. Similarly, different individuals may hold sharply distinct narratives about a global pandemic by drawing on different causal and analogical theories (see Fig. 5 in Section 5).

Despite the ecumenical representational format, narratives are constrained by their functions: They explain and summarize data, facilitate predictions, and motivate and support action. These correspond to the three key processes underlying individual decision-making in CNT, which are intertwined with narratives (Fig. 2).

***Processes.*** *Explanation* makes sense of available data in a unified mental framework by evaluating potential narratives (Section 6). For example, the widow would consider the evidence – time of day, type and duration of noise – to adjudicate the burglar versus noisy cat narratives. Explanation draws on multiple sources of evidence, including prior beliefs, shared narratives, and new observations. Because probabilities are not available under radical uncertainty, heuristics – simple rules relying on small numbers of cues – are used to evaluate narratives, including those exploiting causal, analogical, and temporal structure embedded in narratives. These heuristics are often implemented through affect – which narrative *feels* right.

https://doi.org/10.1017/S0140525X22001157 Published online by Cambridge University Press

**Table 1.** Elements of Conviction Narrative Theory

| | | |
|---|---|---|
| **Context** | Radical uncertainty | Many everyday decisions require beliefs about outcomes that are not finitely enumerable nor their probabilities calculable. |
| | Fuzzy evaluation | Many everyday decisions require trade-offs of values that are incommensurable across choice objects and unstable over time. |
| | Social embeddedness | Beliefs and values are influenced by social context. |
| **Representations** | Narratives | Structured, higher-order mental representations incorporating causal, temporal, analogical, and valence information about agents and events, which serve to explain data, imagine and evaluate possible futures, and motivate action. |
| | Imagined futures | Iconic representations of specific sequences of imagined events generated from a narrative in response to a particular choice being contemplated. |
| | Narrative fragments | Subsets of the elements in a narrative which can be readily communicated. |
| | Shared narratives | Elements of narratives held in common across members of a social group, transmitted through narrative fragments. |
| **Processes** | Explanation | The process of selecting and constructing narratives based on evidence available from the informational, social, and internal cognitive environment, using heuristics and affect. |
| | Simulation | The process of generating imagined futures by projecting a narrative forward. |
| | Affective evaluation | The process of evaluating imagined futures by reacting emotionally to them. |
| | Communication | The processes by which narratives are socially shared, propagating through social networks. |

Description of key aspects of the decision-making context, mental representations, and mental processes invoked by CNT.

**Table 2.** Propositions of Conviction Narrative Theory

| | Proposition |
|---|---|
| Narratives | Narratives are structured, higher-order representations. |
| | Narratives characterize real-world decisions under radical uncertainty. |
| Explanation | We use a suite of explanatory heuristics to evaluate narratives. |
| | Explanatory fit is experienced affectively. |
| Simulation | Imagined futures are simulated by projecting a narrative forward. |
| | Imagined futures are simulated one at a time. |
| Affective evaluation | Affective evaluations of imagined futures motivate choices. |
| | Imagined futures can be appraised on default or ad hoc dimensions. |
| | Emotions are used to manage decisions extended over time. |
| Communication | Shared narratives facilitate social coordination. |
| | Shared narratives shape social learning and evolve. |
| | Shared narratives propagate through social networks. |

These propositions are elaborated in Sections 5–9 with supporting evidence.

The narrative is then used to *simulate* the future (Section 7). Given the burglary narrative, the widow would consider the likely outcome if she were to investigate (being violently attacked), ignore the noise (losing possessions), or equip her investigation with a baseball bat (showing the burglar who's boss). Whereas explanation works by thinking *across* narratives and adopting the most plausible, simulation works by thinking *within* the adopted narrative and imagining the future. We project ourselves into a narrative and imagine what would happen if an action is taken by "running" causation forward. This process generates a representation we term an *imagined future* – a specific sequence of imagined events, represented iconically with a temporal dimension; unlike a narrative, it need not include detailed relational information except an ordered sequence of events. However, imagination has sharp limits – rather than imagining multiple potential futures and "averaging" them, we typically imagine only one future for each choice.

We then *affectively evaluate* that imagined future and take (sustained) action (Section 8). Emotional responses to that future combine beliefs and values. When emotions such as excitement or anxiety are triggered by contemplating an imagined future, we are motivated to approach or avoid that future (Elliot, 2006). The widow imagines an unpleasant future from ignoring the noise, and a more palatable one from a cautious investigation, motivating approach toward the latter future. CNT describes two ways emotions can appraise futures: A *default strategy* based on typical appraisal dimensions used by our affective systems, and an *ad hoc strategy* based on the active goal(s) in our goal hierarchy (Section 8.2).

Emotions are also needed for *maintaining* decisions (Section 8.3). Uncertainty breeds ambivalence, with good arguments for multiple options and the need to sustain decisions over time. When emotions become embedded in narratives, such conviction narratives can manage the incorporation of new information into decision-making while maintaining commitment. As the widow approaches the noise's source, it is natural to feel deeply ambivalent about her choice. Confidence in a stable narrative and imagined future helps to maintain a consistent course of action. When used adaptively, decision-makers incorporate new information from the world into their narratives, creating feedback loops and allowing us to improve repeated or sustained decisions over time.

https://doi.org/10.1017/S0140525X22001157 Published online by Cambridge University Press

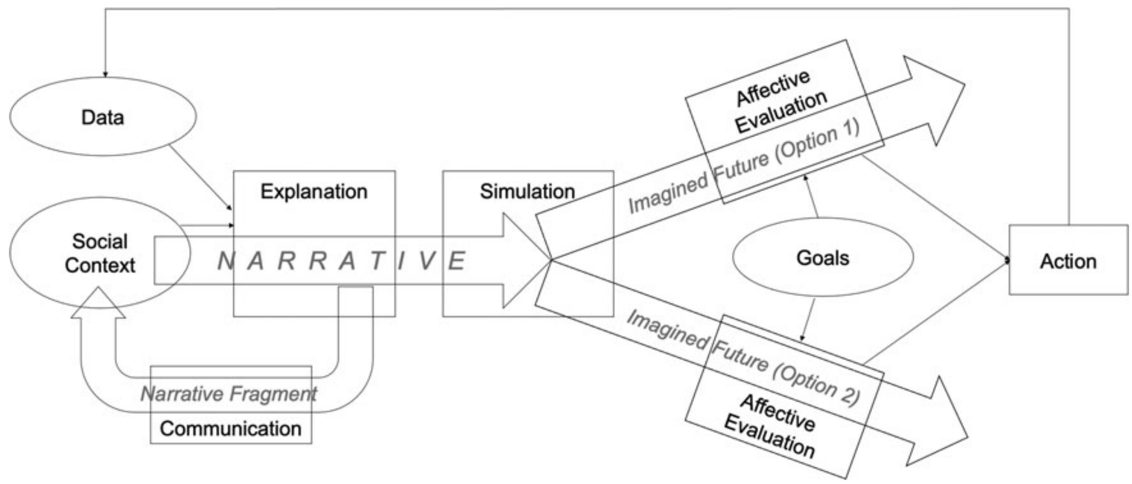


**Figure 2.** Representations and processes in Conviction Narrative Theory.
Narratives, supplied in part by the social environment, are used to explain data. They can be run forward in time to simulate imagined futures, which are then evaluated affectively considering the decision-maker's goals. These appraisals of narratives then govern our choice to approach or avoid those imagined futures. The figure also depicts two feedback loops: Fragments of narratives that are successfully used may be communicated recursively back to the social context, evolving narratives socially, and our actions generate new data that can lead us to update narratives, evolving narratives individually. (*Block arrows depict representations; rectangles depict processes; circles depict sources of beliefs and values, which are inputs to processes via thin arrows*.)

Beyond these backbone operations of choice – explanation, simulation, evaluation – narratives underlie a fourth function: *Communication* (Section 9). Whereas narratives in our definition are mental representations embedded in individual minds, some elements are shared in common across a social group; we refer to this set of common elements as a *shared narrative*. Since these elements are shared only piecemeal (primarily through language), it is these *narrative fragments* that are communicated and which shape and maintain shared narratives. Communication is another way that narratives can participate in feedback loops, now at the collective level: Shared narratives propagate, adapt, and die according to the principles of cultural evolution, permitting learning not only at the individual but at the cultural level. Shared narratives facilitate coordination when used to persuade others and maintain reputation. They propagate when they are catchy enough to be shared, memorable enough to store, and relevant enough to guide decisions.

Nowhere in this proposed process do probabilities or utilities appear; instead, ecologically and cognitively available substitutes play leading roles. In lieu of probabilities to assess narratives, heuristics are used, with narratives arising from the social environment and subjected to cultural evolution; instead of probabilities assigned to imagined futures, the single likeliest imagined future is adopted and evaluated. Rather than utilities assigned to particular outcomes over many dimensions, emotions are felt in response to imagined futures, dependent on the decision-maker's goals.

## 4. Relationship to alternatives

Although we believe our model is the most comprehensive explanation of how and why narratives predominate in decision-making, CNT draws on several related approaches.

### *4.1. Rational approaches*

Bayesian cognitive science models go beyond classical decision theory in showing how probabilities can be calculated and applied to many tasks (Tenenbaum, Kemp, Griffiths, & Goodman, 2011). First, such models specify the hypothesis space. For example, the CEO considering a new blender model might entertain three hypotheses: "We cannot engineer the blender," "We engineer the blender but cannot successfully market it," "The blender expansion succeeds." The reasoner assigns prior probabilities to the hypotheses, evaluates each hypothesis's fit with the data, and combines these using Bayesian inference. For example, the CEO might assign priors of 0.4, 0.35, and 0.25 to these hypotheses. As evidence accumulates – engineers develop a prototype, marketers run focus groups – she considers the likelihood of this evidence under each hypothesis, using Bayes' theorem to combine likelihoods with priors. Despite the simplicity of this toy example, Bayesian cognitive science models make a range of substantive and interesting claims about how people think in a vast array of contexts.

Many critiques have been written of these models (Bowers & Davis, 2012; Jones & Love, 2011; Marcus & Davis, 2013), and we do not endorse all points of criticism. From our perspective, Bayesian approaches potentially work, both in principle and practice, quite well under risk. But their inescapable limitation is the same as for classical decision theory – probabilities cannot be modeled under radical uncertainty, which can interfere with each modeling step:

- **Hypothesis space.** Many problems resist an enumeration of possible outcomes. The CEO has neglected many other possibilities – competitors entering the blender market, engineers generating a prototype no better than competitors', regulators blocking the expansion. This is due to both aleatory uncertainty (some possibilities cannot be imagined even in principle because the future is unknown – the "unknown unknowns") and to generation limits (our finite cognitive capacity to imagine possibilities in open-ended problems).
- **Priors.** It is often unclear even in principle how to rationally set priors. How did the CEO assign a 0.4 probability to her engineering team's failure? Why not 0.2, 0.35, or 0.7? Is it reasonable to use the base rate of engineering team failures given

that this product has never before been designed? This is a specification limit – one cannot non-arbitrarily favor one value over another within a range of plausible values. This not only limits the psychological plausibility of these models, but can be problematic for the models themselves since the specific priors chosen sometimes drive the model's fit (Marcus & Davis, 2013).

- **Likelihoods.** Likelihoods reflect the probability of the evidence conditional on each hypothesis. This raises three problems in realistic contexts. First, information limits: How do we know what evidence is relevant? The CEO may scan the newspaper, consumer research, and marketing reports to grasp the blender market, but will struggle to know which pieces of information bear on her hypotheses. Second, as for priors, specification limits: How do we assign probabilities non-arbitrarily? How is she supposed to estimate the likelihood of particular focus group feedback conditional on the product's future success? Third, capacity limits: With very many pieces of evidence and hypotheses, the amount of calculation rubs up against memory and attention limits.
- **Updating.** Bayesian inference itself quickly approaches capacity limits as the number of hypotheses increases, since priors and likelihoods for each hypothesis must be stored and combined. This is especially problematic for the chains of inference that are common in real-world problems. The probability of successful roll-out depends on the quality of the product, the performance of the marketing team, word-of-mouth, predictions made by analysts and retailers, and countless other factors, often mutually dependent. Moreover, the problems surrounding hypothesis spaces, priors, and likelihoods compound at each step.

Rational approaches – both classical and Bayesian varieties – are nonetheless valuable. They can characterize "small-world" problems, such as risky decisions with enumerable outcomes and probabilities; sometimes provide normative benchmarks for assessing human decisions; provide valuable insights for designing artificial systems (Lake, Ullman, Tenenbaum, & Gershman, 2017); and provide insight at Marr's (1982) computational level by characterizing the goals of a cognitive system. And although probabilistic approaches cannot capture cognition under radical uncertainty, they have inspired some of the boundedly rational approaches discussed next.

### 4.2. Boundedly rational approaches

Both classical and Bayesian approaches have been criticized for their lack of psychological realism, leading to several varieties of *bounded rationality* as amendments. Their core insight is that, although our minds are limited and prone to error, we get quite far with these limited resources: We can be rational within the bounds of our finite minds.

The dominant theoretical style in behavioral economics is surprisingly continuous with neoclassical modeling. Traditional models assume that economic agents are rational, then specify the institutional environment through which the agents interact (e.g., firms and consumers in a competitive market with priced goods) and examine the resulting equilibrium. Behavioral models use the same steps, but tweak the agents to incorporate biases or non-standard preferences, as explained in Section 2.

Like classical approaches, we believe these models produce valuable insights, particularly how small changes to the assumed psychology of economic agents qualify the results of standard models. Yet such models struggle with radical uncertainty and fuzzy evaluation. The same problems that plague classical models with probabilities apply to behavioral models with "decision weights": Such models may capture real psychological biases in how people process probabilities, yet assume probabilities *exist* to be processed. This makes sense for some formal models and laboratory tasks, but not when probabilities do not exist (Section 2.2.1). Likewise, models that stuff the utility function with goodies can capture genuine trends in preferences, but create an illusion of precision when evaluation is fuzzy and options are incommensurable (Section 2.2.2).

Several important principles of bounded rationality, however, do not depend on the intelligibility of optimization:

- **Resource rationality.** In coining the term "bounded rationality," Simon (1957) did not view humans as capriciously irrational, but as managing the best we can given our cognitive and environmental limitations. This approach has been refined in sophisticated models of *resource rationality* (Lieder & Griffiths, 2020), emphasizing the rationality of simplifying strategies such as sampling (Sanborn, Griffiths, & Navarro, 2010). Rationing limited resources is one reason to adopt simplifications – such as narrative thinking – in the face of the calculational difficulties of uncertainty. Equally importantly, probabilistic strategies under uncertainty are not always capable of giving any answer at all.
- **Heuristics.** A heuristic is a fallible-but-useful shortcut. Heuristics are often discussed in contexts where correct answers exist but algorithmic approaches are infeasible or knowledge too limited to provide an optimal answer: Some researchers ("heuristics-and-biases") emphasize the "fallible" part of "fallible-but-useful," and others ("fast-and-frugal") the "but-useful" part (Gigerenzer, 2008; Tversky & Kahneman, 1974). For our purposes, we note that heuristics also may be useful in situations where no objectively correct answer exists, yet some answers are more reasonable than others.

  A parallel debate has raged in normative and descriptive ethics. Utilitarianism emphasizes calculation (Bentham, 1907/1789), positing a duty to maximize social utility. Yet just as Bayesian calculations are often impossible in principle, utilitarian calculations often fail in real-world situations. Aristotle (1999/350 BCE) bemoans the impossibility of a complete theory of ethics, instead urging us to cultivate habit and virtue to do the right thing in particular situations. Indeed, people distinguish between "rational" and "reasonable" behaviors (Grossmann, Eibach, Koyama, & Sahi, 2020), with the former characterized by optimization and abstract universalism, the latter by pragmatism and context-sensitivity (Rawls, 2001; Sibley, 1953). This is likely why descriptive accounts of moral decision-making point to tools such as rules (Greene, Sommerville, Nystrom, Darley, & Cohen, 2001; Kant, 2002/1796; Mikhail, 2007), norms (Nichols, 2002), sacred or protected values (Baron & Spranca, 1997; Tetlock, 2003), and character virtues (Johnson & Ahn, 2021; Uhlmann, Pizarro, & Diermeier, 2015), which often act like heuristics (De Freitas & Johnson, 2018; Sunstein, 2005). For moral decisions, like many everyday choices, often no clearly correct option exists, yet some are more readily justifiable.
- **Ecological rationality.** A crucial point made by some researchers from boundedly rational traditions is that decision-making is adapted to real environments, so seemingly irrational behaviors observed in atypical contexts may be manifestations of more deeply rational – or at least adaptive – behaviors (Todd

https://doi.org/10.1017/S0140525X22001157 Published online by Cambridge University Press

& Gigerenzer, 2007). If most everyday decisions are taken under radical uncertainty, then behaviors that may be adaptive in the real world may manifest as demonstrably suboptimal decisions in the context of risky (often lab-based) contexts.

Narrative approaches to decision-making are compatible with these insights, and can be considered a species of bounded rationality – albeit, at least for CNT, one for which the appropriate benchmark is *reasonableness* rather than *rationality*.

### 4.3. Narrative approaches

Several researchers in both psychology and economics have argued that narratives guide decision-making.

From early days in the heuristics-and-biases tradition, causal thinking was thought to play a privileged role in judgment (Kahneman & Tversky, 1982), such as our ability to use base rates (Ajzen, 1977; Krynski & Tenenbaum, 2007; Tversky & Kahneman, 1980; cf. Barbey & Sloman, 2007; Gigerenzer & Hoffrage, 1995; Koehler, 1996). However, the first decision-making model to consider detailed cognitive mechanisms underlying narrative thought was the Story Model of Pennington and Hastie (1986, 1988, 1992, 1993), most famously applied to juror decisions. In their model, jurors reach verdicts by constructing causal stories and assigning the story to the most appropriate verdict category (e.g., manslaughter, not-guilty). In their studies, participants generated verdicts based on realistic trial evidence. When describing their reasoning, participants overwhelmingly supported their verdicts with causal stories (describing intentions and behaviors) rather than unelaborated lists of facts, with these stories differing greatly across individuals depending on their verdict (Pennington & Hastie, 1986). Manipulating the ease of constructing coherent stories (scrambling evidence order) dramatically shift participants' verdicts (Pennington & Hastie, 1988, 1992). Although the Story Model's legal applications are best-known, it has also been applied to other contexts including economic decisions (Hastie & Pennington, 2000; Mulligan & Hastie, 2005).

Research since this seminal work has developed in two directions. In cognitive science, increasingly sophisticated theories model how people think about networks of causal relationships (Gopnik et al., 2004). For example, Sloman and Hagmayer (2006) argue that people conceptualize their decisions as interventions on a causal network – an idea in sympathy with CNT, wherein choice points in a narrative are opportunities to select among different imagined futures implied by the narrative. A separate but kindred line of work on the Theory of Narrative Thought (Beach, 2010; Beach, Bissell, & Wise, 2016) emphasizes the pervasive role of narratives in memory and cognition, and, like CNT, highlights the importance of narratives for forecasting (Beach, 2020).

A second direction (Abolafia, 2020; Akerlof & Shiller, 2009; Akerlof & Snower, 2016; Shiller, 2019) emphasizes the role of shared narratives in economic outcomes. Shiller argues that when shared narratives go viral, they influence expectations about the future, shaping macroeconomic activity. Shiller's view also provides a powerful role for contagious emotions, especially excitement and panic.

CNT develops these ideas in a third direction: Incorporating ideas about narrative decision-making into a broader framework that elaborates processes and mechanisms, explains how narratives and emotion combine to drive and support action, and accounts for the role of cultural evolution of narratives to render decision-making adaptive even under radical uncertainty. We see CNT as complementing rather than contradicting these perspectives, developing these approaches to the next stage in their evolution.

## 5. Narratives

Prior work has not coalesced around a single definition of "narrative," much less a single notion of representation. For example, Beach (2010) defines "narrative" as "…a rich mixture of memories, of visual, auditory, and other cognitive images, all laced together by emotions to form a mixture that far surpasses mere words and visual images in their ability to capture context and meaning," while Shiller (2019) follows the Oxford English Dictionary in defining it as "…a story or representation used to give an explanatory or justificatory account of a society, period, etc." (quoted in Shiller, p. xi). Meanwhile, Pennington and Hastie (1992) say that stories "…could be described as a causal chain of events in which events are connected by causal relationships of necessity and sufficiency…." The hierarchical structure of stories – for instance, that events can be grouped into higher-order episodes – is also often noted as a common feature (Abbott, Black, & Smith, 1985; Pennington & Hastie, 1992). Across these conceptions, causation is central but not the only hallmark of narratives – they provide meaning by explaining events (Graesser, Singer, & Trabasso, 1994; Mandler & Johnson, 1977; Rumelhart, 1975).

In our view, ordinary causal models (Pearl, 2000; Sloman, 2005; Spirtes, Glymour, & Scheines, 1993) are a crucial starting point, yet not quite up to the task of representing narratives. (That said, some progress has been made toward formalizing some economic narratives in this way; Eliaz & Spiegler, 2018.) For our purposes, causal models have two shortcomings: They do not represent some information – such as analogies and valence – that will prove crucial to narrative thinking; and operations over causal models are usually assumed to be probabilistic – a non-starter under radical uncertainty.

We define narratives as *structured, higher-order mental representations incorporating causal, temporal, analogical, and valence information about agents and events, which serve to explain data, imagine and evaluate possible futures, and motivate and support action over time*. No doubt, this definition itself requires some explanation.

### 5.1. Narratives are structured, higher-order representations

In a *structured* mental representation, relations are represented among the objects it represents. For example, a feature list is an unstructured representation of a category, whereas a propositional, sentence-like representation with explicit predication (i.e., assigning attributes to specific elements and specifying relations among elements) is highly structured. Similarly, causal models are highly structured, as are representations of categories whose features are related to one another through analogy (Gentner, 1983).

Narratives may have been difficult to pin down in past work because they are structured, like causal models, but contain richer information that is not typically represented in causal models. Specifically, we argue that narratives can represent causal, temporal, analogical, and valence structure. Complicating things further, not all types of structure are necessarily invoked in a given narrative. Narratives are *higher-order representations* that flexibly include lower-order representations.

https://doi.org/10.1017/S0140525X22001157 Published online by Cambridge University Press

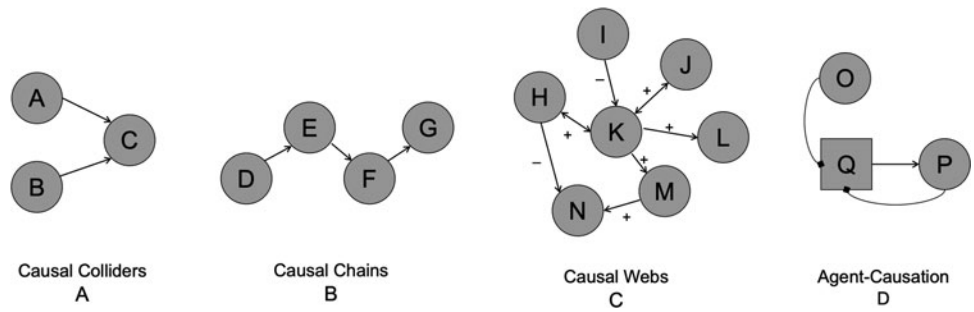


**Figure 3.** Common causal structures in narratives.
In panel *A* (a *causal collider*), multiple potential causes (A or B) could explain an event (C); a typical inference problem would be to evaluate A and B as potential explanations of observation C, which may in turn license other inferences about effects of A or B (not depicted here). In panel *B* (a *causal chain*), a sequence of causally related events (D, E, F, G) is posited; typical inference problems would be to evaluate whether the overall sequence of events is plausible, or whether an intermediate event (E) is plausible given that the other events (D, F, G) were observed. In panel *C* (a causal web), many event types (H–N) are thought to be related to one another, with some relationships positive and others negative, and some bidirectional; typical inference problems would be to evaluate the plausibility of individual links or to infer the value of one variable from the others. In panel *D* (*agent-causation*), an agent (Q) considers taking an action (P), based partly on reasons (O) and their judgment of the action itself (P); typical inference problems would be to predict the agent's action based on the available reasons, or infer the agent's reasons based on their actions. [*Circles and squares depict events and agents, respectively; straight arrows depict causal relationships, which could be unidirectional or bidirectional, positive (default or with a "+" sign) or negative (with a "–" sign); curved, diamond-tipped arrows depict reasons. For causation among events and agents, but not event-types (panels A, B, and D), left–right orientation depicts temporal order.*]

#### 5.1.1. Causal structure

Narratives represent at least two types of causal relationships: *event-causation* and *agent-causation*. Event-causation can further be sub-divided based on the kind of event (individuals or categories) and how the events are connected (e.g., colliders, chains, or webs).

*Event-causation* refers to dependency relationships between either individual events (*the Central Bank lowered interest rates, causing investment to increase*) or event categories (*lower interest rates cause increased investment*). Both kinds of event-causal relations are important since narratives incorporate information both about individual events (e.g., the course of my marriage) and event categories (e.g., how relationships work generally), including analogical links between these knowledge types. We are agnostic about the representational format of event-causation, and indeed these representations may involve aspects of networks, icons, and schemas (Johnson & Ahn, 2017). For familiarity, the diagrams we use to depict narratives are elaborated from causal network conventions (Figs. 3–6).

Patterns of event-causation also vary in their topology and inference patterns (Johnson, 2019). Three common types of causal patterns in narrative contexts are colliders, chains, and webs (Fig. 3A–C).

In a *causal collider* (Fig. 3A), we can observe evidence and seek an explanation for it, which may in turn generate further predictions. For example, if a central banker makes some statement, this licenses inferences about the banker's intention, which may yield predictions about the bank's future policies.

In a *causal chain* (Fig. 3B), a sequence of events is causally and temporally ordered. Mrs O'Leary went to milk her cow; the cow objected and kicked a lantern; the lantern started a fire; and so began the Great Chicago Fire (supposedly). A mysterious individual invents blockchain technology; it fills an important economic niche; it gains value; it becomes widely adopted. We can think about the plausibility of these sequences overall, fill in missing events from the chains, and predict what will happen next.

In a *causal web* (Fig. 3C), we ask how a set of variables relate to one another – an intuitive theory (Shtulman, 2017). Some intuitive theories probably have innate components (Carey, 2009a), but many decision-relevant intuitive theories are learned. For example, investors must have mental models of how macroeconomic variables such as exchange rates, inflation, unemployment, economic growth, and interest rates are linked (Leiser & Shemesh, 2018), and voters have intuitions about trade, money, and profits (Baron & Kemp, 2004; Bhattacharjee, Dana, & Baron, 2017; Johnson, Zhang, & Keil, 2018, 2019). These intuitions likely drive much political and economic behavior, yet differ strikingly from economists' consensus (Caplan, 2007; Leiser & Shemesh, 2018). This may reflect both divergences between the modern world and evolved intuitions (Boyer & Petersen, 2018), and the difficulty of correctly extracting causal structure from causal systems with more than a few variables (Steyvers, Tenenbaum, Wagenmakers, & Blum, 2003).

Narratives often center around the intentional actions of human agents, and such *agent-causation* appears to be a very different way of thinking about causation. Rather than representing *events* as causing one another (e.g., Alan Greenspan's forming the intention to decrease interest rates caused interest rates to decrease), people sometimes appear to think of *agents* as causing events directly (e.g., Alan Greenspan caused interest rates to decrease). This reflects the intuitive notion that agents have free will; that our choices, when construed as agent-causes rather than event-causes, are themselves uncaused (Hagmayer & Sloman, 2009; Nichols & Knobe, 2007). Agents can act for *reasons* (Malle, 1999); they make intentional choices based on their beliefs and desires which are assumed to be rational (Gergely & Csibra, 2003; Jara-Ettinger, Gweon, Schulz, & Tenenbaum, 2016; Johnson & Rips, 2015). Complicating things further, reasons are often anticipations of the likely effects of one's action.

#### 5.1.2. Temporal structure

Narratives often, but not always, include temporal information about the order, duration, and hierarchical structure of events. For example, causal chains are necessarily ordered sequences of events because causes occur before effects. At the opposite extreme, temporal order often is lacking entirely from causal webs that depict causality among event *categories* rather than individual events. That said, people can track the order, delay, and part–whole structure of causally related events and use these different types of temporal information to disambiguate causal

https://doi.org/10.1017/S0140525X22001157 Published online by Cambridge University Press

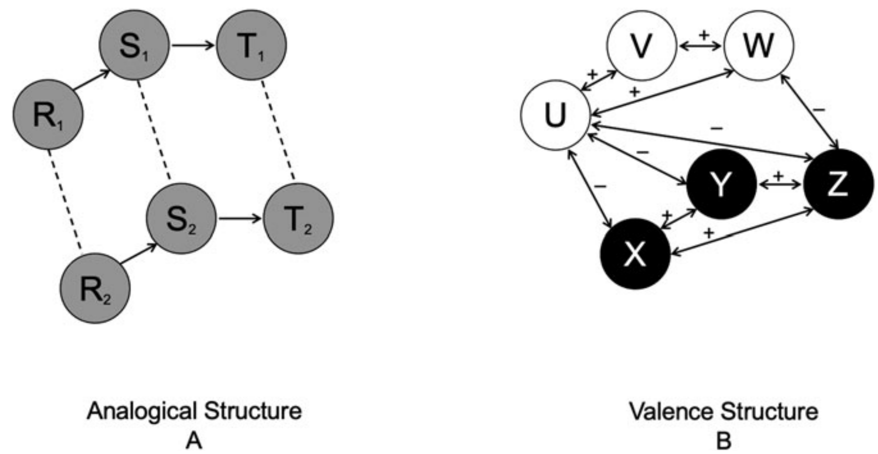


**Figure 4.** Analogical, valence, and causal structure.
In panel *A* (*analogical structure*), one causal chain ($R_1$, $S_1$, $T_1$) is analogized to another ($R_2$, $S_2$, $T_2$); typical inference patterns would be to reason from a known sequence ($R_1$, $S_1$, $T_1$) of specific events or schematized depiction of a general causal mechanism to infer the causal–temporal order of a new sequence ($R_2$, $S_2$, $T_2$) or to infer missing events ($T_2$) given that all other events are observed. In panel *B* (*valence structure*), positive event types (U, V, W) are seen as bidirectionally and positively related to each other, negative event types (X, Y, Z) are seen as bidirectionally and positively related to each other, whereas negative and positive events are seen as bidirectionally and *negatively* related to each other (Leiser & Aroch, 2009). (*Dashed lines represent analogical correspondences; white and black circles represent "good" and "bad" events or event types, respectively.*)

structures (Lagnado & Sloman, 2006; Rottman & Keil, 2012). For example, sequences of events often are segmented into higher-order episodes, each containing lower-level sub-events (Zacks & Tversky, 2001). This part–whole organization affects causal representations, with higher-level events believed to be both causes and effects of other higher-level events, and low-level events from one high-level cluster believed to affect only low-level events from that same cluster (Johnson & Keil, 2014).

#### *5.1.3. Analogical structure*

Both the power and peril of narrative thinking compound when people perform inference not only by causal thinking within a single domain, but across different domains through analogies. Structure-mapping theory (Gentner, 1983) is a model of how people select analogies and use them to make inferences, emphasizing that matches in the *relationships* within a domain make it a good or bad analogy for another domain. Thus, analogy is especially powerful when combined with other relational systems such as causal systems (Holyoak, Lee, & Lu, 2010) (Fig. 4A).

Analogies are important to narratives for at least two reasons. First, they allow us to use familiar domains to make sense – if imperfectly – of less familiar domains. For example, people have highly impoverished mental models of central banking, but more detailed mental models of cars. In a car, stepping on the gas pedal causes more gasoline to enter the engine, increasing the car's speed. People often use this analogy for discussing and understanding central banking; the central bank prints more money, causing more money to enter the economy, causing the economy to go faster. Second, abstract and gist-like representations apply to a broader set of future situations, particularly when making decisions about the distant future (Schacter & Addis, 2007; Trope & Liberman, 2003). Thus, forming analogical links among specific past events and, ultimately, between specific events and more abstract event categories is crucial for generating generalizable knowledge. This is how our representational system incorporates some aspects of narratives' hierarchical structure.

#### *5.1.4. Valence structure*

Stories involve good guys and bad guys, goals being achieved or objectives thwarted. Information about norms (right or wrong) and valence (good or bad) is processed rapidly and automatically (Moors & De Houwer, 2010) and influences thinking in many domains including causation (Knobe, 2010). For example, people are likelier to identify norm-violations as causes and non-norm-violations as non-causes (Kominsky, Phillips, Gerstenberg, Lagnado, & Knobe, 2015), reason differently about the potency of good versus bad causes (Sussman & Oppenheimer, 2020), and tend to think that good events cause other good events (LeBoeuf & Norton, 2012). Macroeconomic understanding is dominated by a "good begets good" heuristic (Leiser & Aroch, 2009), wherein "bad" events (inflation, unemployment, stagnation, inequality) are thought to be inter-related and negatively related to "good" events (price stability, full employment, economic growth, equality) (Fig. 4B). In reality, the opposite often holds.

#### *5.1.5. Coherence principles*

Given their rich representational capacities, *coherence principles* are needed to constrain narratives' vast possibility space (determining *which* narratives to entertain) and draw inferences about missing information (filling in details *within* a narrative). For example, Thagard (1989) develops a theory of how explanations and evidence cohere, Gentner (1983) presents a theory of how analogical correspondences are drawn, and Rottman and Hastie (2014) summarize evidence about how people draw inferences on causal networks. In addition to principles governing each type of lower-level representation individually, we suggest three principles as starting points for how different types of lower-level representations cohere: (i) Causal, temporal, and valence structures are preserved across analogies; (ii) Causes occur before their effects; (iii) Causal relationships between agents and events with the same valence status ("good" or "bad") are positive, whereas they are negative for links between "good" and "bad" events.

#### *5.1.6. What narratives are not*

Narratives are a flexible representational format, but they are not *infinitely* flexible. We (tentatively) suggest the following test for whether a representation is a narrative: It must (i) represent causal, temporal, analogical, or valence information, and (ii) for any of these it does *not* represent, it must be possible to incorporate such information.

This distinguishes narratives from several other kinds of representations, including probabilities, spatial maps, associative networks, images, categories, and logical relations: Such formats do not necessarily include any of the four structured information types. (However, elements of narratives may be linked to such other representations in memory. Indeed, this may be required

https://doi.org/10.1017/S0140525X22001157 Published online by Cambridge University Press

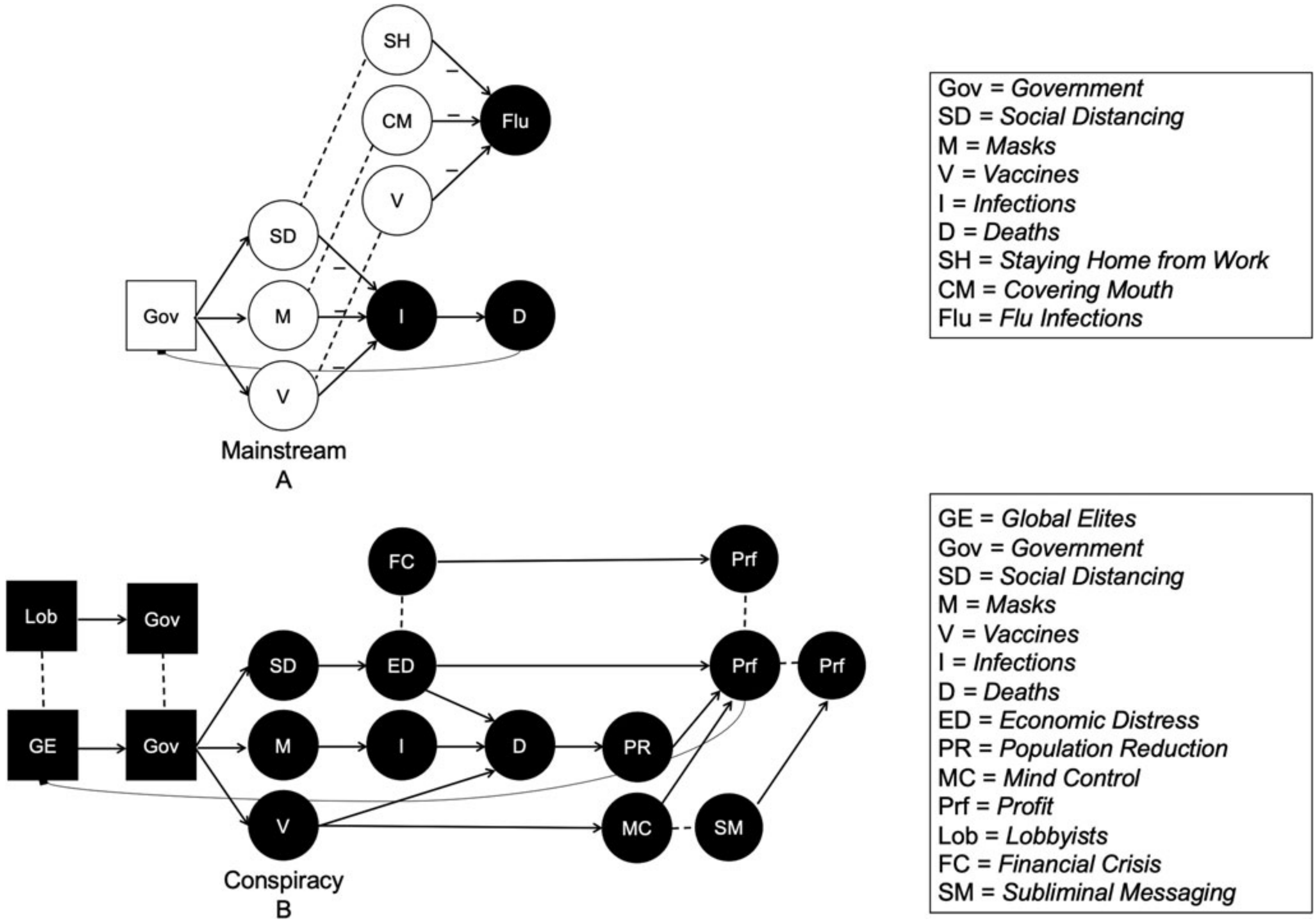


**Figure 5.** Possible narratives around a global pandemic.
Panel *A* depicts one possible individual's narrative around a global pandemic, which aligns largely with the mainstream view. Infections and deaths (which are bad) are negatively related to interventions such as social distancing, masks, and vaccines, which are themselves results of government action. The government chose these actions for the reason that it would have a preventative effect on deaths. The causal links between each intervention and infection is supported by an analogy to other diseases, such as influenza (i.e., staying home from work, covering one's mouth when coughing, and vaccines all help to prevent flu infections). Panel *B* depicts one possible conspiratorial narrative around a pandemic. In this narrative, global elites control the government, and are acting so as to increase their profits, which can be accomplished by several channels including economic distress, population reduction, and mind control. These causal links to profitability are supported by their own analogies (e.g., the global financial crisis and subliminal messaging being ways that bankers, corporations, and other elites are thought to increase their profits), as is the idea that the government is captured by unelected elites such as lobbyists for big business. In this narrative, social distancing has little effect on the spread of disease but a strong link to intentional economic distress; masks and vaccines increase infection and death rather than preventing it. For this reason, interventions that are seen as good in the mainstream narrative (because they have a preventive relationship with death) are seen as bad in the conspiracy narrative. These hypothetical narratives will be supported by different social and informational environments, yield conflicting forecasts about the future, and motivate distinctive actions.

for narrative simulation to generate iconic representations of imagined futures.)

Figure 5 provides additional examples of possible narratives that might underlie decision-making in the context of a global pandemic, as in one of our running examples.

### 5.2. Narratives characterize real-world decisions under radical uncertainty

Lab experiments are ill-suited for testing the prevalence of narrative thinking in everyday decision-making. Thus, we bring linguistic and qualitative data to bear.

#### 5.2.1. Linguistic data about macroeconomic narratives

Shiller (2019) uses time-series data from Google N-Grams to track language linked to particular shared narratives. Shiller emphasizes that "viral" narratives, even if false, can affect macroeconomic events. We would add that shared narratives held true within one's own social network (rather than those held true somewhere else) may be the only way for many to make sense of complex macroeconomic causation.

The shared narrative that labor-saving machinery creates unemployment is perennial. Shiller traces it from Aristotle, through worker riots during the Industrial Revolution, to economic depressions, to present-day concerns about artificial intelligence displacing humans. (Most economists disagree: Since machines increase productivity, wages rise and labor is redeployed to higher-valued uses.) Shiller traces the frequency of "labor-saving machinery" in books from 1800 to 2008, with the term peaking during the 1870s depression and again in the lead-up to the Great Depression, with the new term "technological unemployment" reaching epidemic proportions throughout the 1930s. Plausibly, the fear produced by such narratives exacerbated the underlying problems causing the Depression. Shiller notes that

https://doi.org/10.1017/S0140525X22001157 Published online by Cambridge University Press

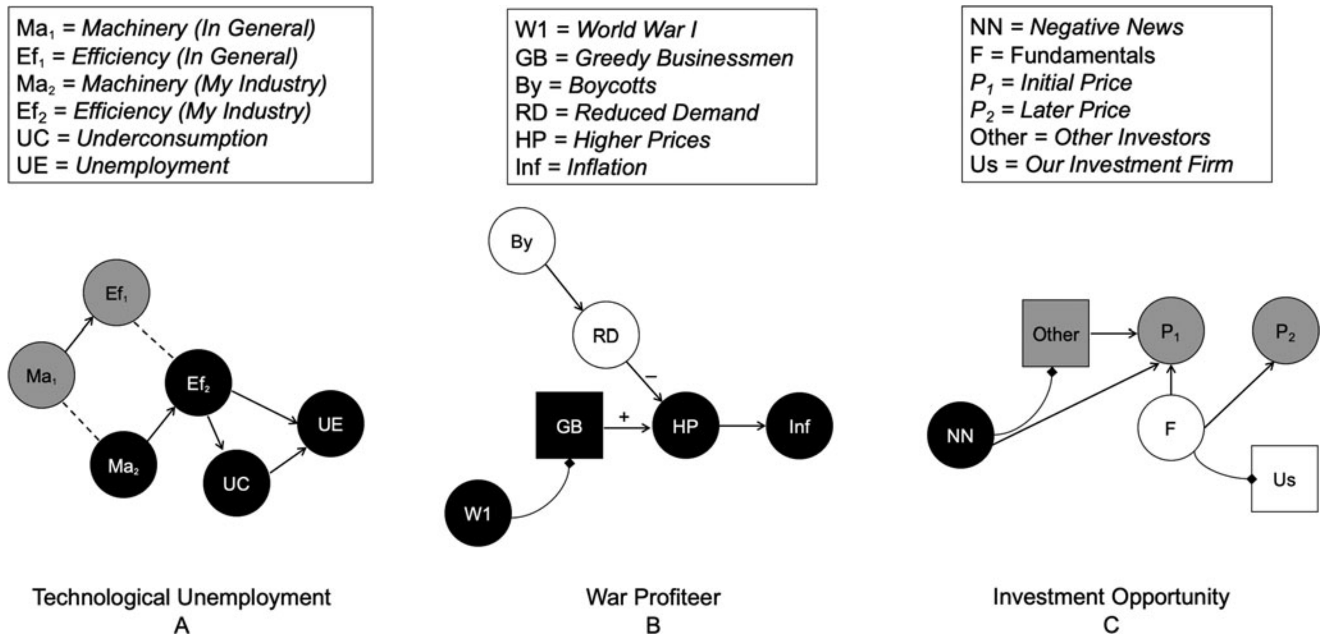


**Figure 6.** Economic narratives from linguistic and interview data.
Panels *A*–*C* depict simplified versions of three narratives drawn from Shiller's (2019) linguistic studies of viral economic narratives and Tuckett's (2011) interview studies of money managers. In panel *A*, a generic causal mechanism of machinery generally leading to increased efficiency ($Ma_1$ and $Ef_1$) is analogized to machinery in one's particular industry leading to increased efficiency in that industry ($Ma_2$ and $Ef_2$). Efficiency is thought to cause unemployment (*UN*) directly by displacing human workers and indirectly through underconsumption (*UC*). Because unemployment is seen as bad, all other variables in the causal chain are inferred to be bad too. In panel *B*, greedy businessmen (*GB*) are inspired by the opportunity of World War I (*W1*) to increase prices (*HP*), which leads to inflation (*Inf*). A boycott (*By*) is thought to reduce demand (*RD*), which would in turn push prices back down (negative effect on *HP*). Since inflation and the greedy businessmen who cause it are bad, the countervailing boycott chain is perceived as good. In panel *C*, negative news about a company (*NN*) is thought to affect its stock price at an initial time ($P_1$), but only the company's fundamentals (*F*) affect its stock price later ($P_2$). Other investors (*Other*) are less observant and only act based on the negative news, but our investment firm (*Us*) is more observant and sees the fundamentals, creating a profit opportunity.

this dovetails with the then-popular folk theory – never accepted by economists – that machines would produce such plentiful products that we could never consume it all, generating unemployment. Accordingly, the term "underconsumption" skyrockets during the Depression. A simplified version of this narrative can be seen in Figure 6A.

Another of Shiller's examples concerns boycotts following World War I (Fig. 6B). The US dollar experienced 100% inflation after the war, contributing to an anti-business shared narrative, with mentions of "profiteer" in newspapers peaking at the start of the subsequent depression. According to this narrative, businesses were raising prices to achieve "excess profits" during the war, explaining the inflation. (Although economists now reject this view of inflation, similar narratives were proposed by some American politicians during the inflation episode in 2022.) Protests ensued, resulting in boycotts on the theory that if consumers did not buy products beyond minimum necessities, the drop in demand would force prices back to "normalcy." Although prices never declined to pre-war levels, deflation did indeed result in the 1920–1921 depression.

### 5.2.2. Interview data about microeconomic decision-making

In the same spirit, but using a very different method, Chong and Tuckett (2015) and Tuckett (2011, 2012) interviewed 52 highly experienced money managers in 2007 and 2011, gathering accounts of decisions to buy or sell securities, using a standardized non-schedule interview approach (Richardson, Dohrenwend, & Klein, 1965; Tuckett, Boulton, & Olson, 1985).

These accounts consistently invoked narratives. For example, consider one of the respondents selected at random for detailed presentation in Tuckett (2011, p. 33). When interviewed, he directed a team of 20 and was personally responsible for allocating stocks into a $35 billion portfolio. His task was to try "to pierce through the smoke and emotion" surrounding market moves and "be contrary to the consensus notion of 'let's wait for the smoke to clear'." "I mean the problem with that philosophy" as he put it, is that "if you wait for everything to be clear you will miss most of the money to be made." "Once everything's clear…it's easy, right?"

He described one stock his firm had bought the previous year (Fig. 6C), which had been experiencing issues with a main supplier, leading to very negative news. His team, however, "kicked the tyres and did a lot of work" to conclude that the situation was not so dire as widely perceived, taking a large stake that rose over 50% within weeks. After exiting their position, there were more negative news items, involving a very large shareholder selling the company's stock, causing apprehension in the market and price drop; they again concluded that the company was undervalued and "re-established the position." "It was somewhat controversial…. It was not easy going against consensus sentiment but that's…what distinguishes us." It worked out.[1]

Respondents' narratives coalesced around common themes (Tuckett, 2011, p. 89). For instance, among the 165 "Buy" narratives from the 2007 interviews, themes included spotting some attractive features through the respondent's exceptional ability or effort (45% prevalence as rated by independent content-coders); the company/sector offering exceptional opportunities (39%); limits to downward surprise (27%); management as proven and reliable (26%); successful management of the respondent's own emotions (11%); and (temporary) monopoly or market power (10%).

https://doi.org/10.1017/S0140525X22001157 Published online by Cambridge University Press

In addition to providing examples of narrative thinking, these interviews illustrate five features of the radically uncertain decision context money managers face, which suggest why narratives are useful (Tuckett, 2011, pp. 50–54). First, they acted in situations where they could only speculate about the outcome of their actions and doing nothing *was* an action. Second, the financial assets they traded had unknowable future values. Given their mandates to try to outperform each other, they sought opportunities that they thought wrongly priced. To establish mispricing, they imagined how various stories might influence future income streams of a firm and how others would react in those situations. The issue, in the several hundred decisions analyzed, was always the same: How to know and create confidence in their particular story about future prices. Third, the available data to help them form price expectations were effectively infinite – a massive range of public and privately available information, some of questionable provenance, from countless sources in numerous languages. Fourth, they made decisions in a social context. Most respondents talked with or explored their views of the future with others or had to justify their decisions (and reputations) to others. Fifth, decisions were never final and time horizons were always part of the context. For how long would prices go lower before rising? Had prices reached their peak? The money managers' decisions had many of the features depicted in Section 1, which they managed using narratives.

Interviews do not provide conclusive evidence about cognitive processing, but do underscore two key points – the respondents' decision context is characterized by radical uncertainty, and respondents frequently invoke narratives in their reflections on these choices. Plausibly, narratives may have played some role in many of these choices, but finer-grained experimental evidence is needed to understand this role more precisely.

## 6. Explanation

Good decisions about the future often depend on how we understand the present. *Explanation* is how we create these understandings – how narratives are constructed and evaluated.

We have a fundamental motivation to make sense of things (Chater & Loewenstein, 2016) – a drive that pervades mental life. The world is not perceived directly, but must be interpreted (Fodor & Pylyshyn, 1981; Gregory, 1970; Von Helmholtz, 2005/1867). Light hits our two-dimensional retinas, and perception allows us to infer from this a three-dimensional world – to make sense of data under uncertainty. Much cognitive processing has a similar logic: Assembling relevant facts into useful models that can guide predictions and choices. Categories license predictions about individuals, using an object's observed features (evidence) to determine the appropriate category (hypothesis) (Murphy & Medin, 1985). Causal cognition explains events (evidence) in terms of hidden causes (hypotheses) (Lombrozo, 2016). Our memories use scattered strands of remembrance (evidence) to piece together a coherent story about what happened (hypothesis) (Bartlett, 1932; Johnson & Sherman, 1990). Theory-of-mind uses others' observed behavior (evidence) to infer mental states (hypotheses) (Gergely & Csibra, 2003).

Bayesian approaches to explanatory reasoning are popular in both cognitive science and philosophy of science. This approach conceptualizes the phenomenon to be explained (*explanandum*) as evidence and potential explanations as hypotheses to be evaluated using Bayesian inference. Thus, rational explanation requires the reasoner to evaluate the prior probability of each explanation $[P(H)]$ and the fit of the evidence with the explanation $[P(E|H)]$.

In two important respects, this is a successful theory. First, it uses the same mental machinery to understand many explanatory inferences. There is a common logical structure to explanatory inferences such as causal reasoning, theory-of-mind, and categorization; rather than invoking entirely separate mental mechanisms, it assumes common computational mechanisms. This is in keeping with the empirical evidence as well as theoretically parsimonious. Second, it accounts for why people are reasonably adept at explaining events. Enterprises such as science, technology, and commerce – not to mention everyday activities such a social interaction – depend on explanatory processes; to the extent we are adept at these activities, we are necessarily adept at explaining things.

Yet this approach cannot be quite right. For the reasons above (Section 2.2.1), such calculations are often impossible in *principle*, much less for flesh-and-blood humans, given aleatory and epistemic limits on probabilities. Bayesian accounts can try to avoid this problem by claiming agnosticism about the actual processes used to reach the outputs of Bayesian theories (retreating to the computational level) or by invoking approximation mechanisms that do not require the full probabilistic machinery. Despite our sympathy with both approaches, the theoretical problems can only be fully resolved if these strategies avoid invoking probabilities altogether under radical uncertainty.

Even eschewing probabilities altogether – as in some Bayesian sampling approaches (Sanborn & Chater, 2016) – cannot avoid a further problem: A dizzying array of empirical anomalies relative to Bayesian predictions. For instance, people often favor explanations that are simpler than merited by the evidence (Lombrozo, 2007), while in other contexts favoring overly complex explanations compared to rational models (Johnson, Jin, & Keil, 2014). Although people reasonably prefer explanations that explain more observed evidence (Read & Marcus-Newhall, 1993), they also prefer explanations that make no unverified predictions (Khemlani, Sussman, & Oppenheimer, 2011), contradicting Bayes' theorem. Yet these anomalies are not random: They point to systematic principles we use to evaluate explanations in the absence of precise probabilistic reasoning.

### *6.1. We use a suite of explanatory heuristics to evaluate narratives*

Successful explanation under radical uncertainty requires strategies that circumvent probabilistic reasoning, but instead exploit other forms of structure available in a given situation. People use a variety of heuristics and strategies satisfying this criterion (e.g., Horne, Muradoglu, & Cimpian, 2019; Lombrozo, 2016; Zemla, Sloman, Bechlivanidis, & Lagnado, 2017). These strategies are useful, not because they are infallible or optimal (which may not even be meaningful under radical uncertainty), but because they do not require explicit probabilistic reasoning, yet exploit regularities that often make these strategies broadly truth-tracking.

Above, we highlighted how temporal, analogical, and valence information are incorporated into narrative representations. Heuristics exploiting such information are valuable because they capitalize on regularities we naturally attend to, such as event structure and causal mechanisms, while powerfully constraining which explanations are deemed plausible (Einhorn & Hogarth, 1986). For example, people use temporal order (Bramley, Gerstenberg, Mayrhofer, & Lagnado, 2018), the delay between cause and effect (Buehner & May, 2002), and the part–whole structure of events (Johnson & Keil, 2014) to disambiguate causal directionality, determine which events are causally relevant, and assign causal responsibility. Such strategies are adaptive because they exploit regularities in the environment that are less susceptible to the problems of radical uncertainty, even if they are not infallible.

Likewise, analogies allow us to extend hard-won knowledge of one domain into another. Much of our causal knowledge appears to be stored in the format of stereotyped causal schemas (Johnson & Ahn, 2015, 2017). We can analogically reason from a known story (e.g., my cousin's marriage; the Spanish Flu) to the current situation (e.g., my marriage; COVID); or at a more abstract level, we reason from known causal mechanisms when evaluating an unfamiliar domain, as when we compare the economy to a stalled car or a bureaucratized corporation to an arthritic giant. The notion that "good begets good" (Leiser & Aroch, 2009) can be thought of a highly generalized causal schema, with narratives failing to match this schema (e.g., actions that decrease inflation are likely to increase unemployment) fighting an uphill battle for plausibility. In other cases, we have domain-specific expectations (Johnston, Sheskin, Johnson, & Keil, 2018), such as the belief that physical causation follows more linear causal pathways compared to the more web-like structure of social causation (Strickland, Silver, & Keil, 2016).

Other heuristics derive from causal structure itself, accounting for the anomalies above. People often substitute the vague and challenging question of an explanation's prior probability for the more straightforward question of an explanation's simplicity (Lombrozo, 2007). Yet, because simpler explanations often explain less of the data, people also use an opponent *complexity heuristic* to estimate an explanation's fit to the data or Bayesian likelihood (Johnson, Valenti, & Keil, 2019). Likewise, people prioritize explanations that account for a wider range of observations (Read & Marcus-Newhall, 1993) and attempt to make inferences about evidence that has not actually been observed (Johnson, Kim, & Keil, 2016a; Johnson, Rajeev-Kumar, & Keil, 2016b; Khemlani et al., 2011).

Explanatory heuristics are used widely across cognition – for problems such as categorization, causal attribution, theory-of-mind, stereotyping, and even some visual tasks (Johnson, 2016; Johnson et al., 2014; Johnson, Merchant, & Keil, 2015a; Johnson, Rajeev-Kumar, & Keil, 2015b; Johnson et al., 2016a, 2016b; Sussman, Khemlani, & Oppenheimer, 2014) and emerge early in development (Bonawitz & Lombrozo, 2012; Cimpian & Steinberg, 2014; Johnston, Johnson, Koven, & Keil, 2017). They are also linked to action: People favor explanations that license task-relevant and high-utility actions, and which highlight stable causal relationships likely to apply across contexts (Johnson et al., 2015b; Vasilyeva, Wilkenfeld, & Lombrozo, 2017, 2018). Explanatory heuristics, though imperfect, aid people in circumventing specification limits and information limits (Section 2.2.1).

### 6.2. Explanatory fit is experienced affectively

Although these heuristics may function to push us toward more useful or probable portions of the hypothesis space, their phenomenology is often more affective than cognitive. Emotions rapidly summarize information not readily available to consciousness (Rolls, 2014; Rozin & Fallon, 1987; Todorov, 2008). Because emotions have an intelligence of their own (Nussbaum, 2001), we make inferences from them (Cushman, 2020; Schwarz, 1990) and often rely on "gut" feelings to assess situations (Klein, 1998). This can be a broadly rational strategy despite leading to some mistakes.

In the case of explanations, we feel "satisfied" when we achieve a sense of understanding (Gopnik, 1998; Lipton, 2004). Despite the adaptive basis of many of these heuristics, we do not *feel* like we are performing rational computations when using them. Instead, good explanations often are accompanied by positive emotions or even aesthetic beauty (Johnson & Steinerberger, 2019), as when scientists and mathematicians claim to prioritize beauty in constructing their theories. Conversely, explanations that conflict with prior beliefs produce cognitive dissonance and may therefore be rejected (Festinger, 1962).

Although heuristics often lead to error (Kahneman, 2002), they can be adaptive in solving problems given cognitive and environmental limits (Gigerenzer & Goldstein, 1996; Simon, 1955). Good thing, too: Under radical uncertainty, they often are all we have.

## 7. Simulation

Having selected a narrative through explanation, we project the narrative into the future through simulation. This is why sense-making and imagination are linked: We make sense of the past to imagine the future.

### 7.1. Imagined futures are simulated by projecting a narrative forward

The brain mechanisms involved in prospective thought about the future overlap with those used for episodic memory about the past (Schacter, Addis, & Buckner, 2008) and may even be subsystems of a broader *mental time travel* faculty (Suddendorf & Corballis, 1997, 2007). This is consistent with our view that the same representations – narratives – underlie explanations of the past and simulations of the future (Aronowitz & Lombrozo, 2020). Moreover, simulation can rely on step-by-step reasoning using causal mechanisms. For example, when shown a set of interlocked gears and asked which direction one gear will turn given that another gear is turned, people solve this problem by mentally turning one gear at a time (Hegarty, 2004). Likewise, thoughts about how reality might be different – counterfactuals – operate over causal structures (Rips, 2010; Sloman, 2005) and are central to imagination (Markman, Gavanski, Sherman, & McMullen, 1993).

However, less work has looked at how particular features of narratives manifest in simulations or how these simulations manifest in choices. Our recent research program has studied the role of narrative-based simulation in financial decision-making.

Our strategy relies on the idea that cognitive processes have specific signatures associated with their limitations or biases. To use an example from a different area (Carey, 2009b), a signature of analog magnitude representations is that the ability to discriminate two magnitudes is proportional to their ratio. Since discriminations for large (but not small) numbers have this property, this implies that representations of large numbers are analog.

Analogously, we look for signature limitations associated with using narrative representations for predictions and decisions under uncertainty. Since narratives can incorporate causal, valence, temporal, and analogical structure (Section 5.1), we designed experiments to examine whether introducing these structural features into forecasting problems produces signature biases relative to standard financial theory. Although any one study does not necessarily implicate narratives, their combination

https://doi.org/10.1017/S0140525X22001157 Published online by Cambridge University Press

triangulates on the conclusion that people project narratives to simulate the future.

#### 7.1.1. Causal structure: internal and external attributions

Narratives contain causal structure which provides explanations. Therefore, if people use narratives to forecast the future, causal explanations should affect future forecasts. To test this, participants read about companies that had recently changed in stock price (Johnson, Matiashvili, & Tuckett, 2019a). According to financial theory, the reason for the price change is irrelevant to predicting future prices. For example, when a CEO retires, the firm's stock price often declines. But this decline occurs when the CEO's retirement becomes publicly known, after which it does not reliably produce further declines. Nonetheless, if people use stories to predict the future, it should be irresistible to look for *why* the price changed and use these inferred causes to predict further price changes.

In one study, we compared three explanation types for a price change – no explanation, an internal attribution (relating to skill or quality, e.g., an ill-considered management change), or an external attribution (relating to factors outside the company's control, e.g., a market crash). Participants predicted more extreme future trends when an explanation was given rather than not given, and more extreme trends when the explanation was internal rather than external. This suggests that people look for causal stories to account for events and predict future ones, the most compelling stories invoking internal or inherent features (Cimpian & Steinberg, 2014).

Financial markets are volatile: Prices often shift with little apparent cause (Shiller, 1981). Might people nonetheless supply causes of price changes by default, affecting downstream predictions? To find out, we compared the no-explanation and internal-explanation conditions to a noise condition, in which the price change was explained as random. Several conclusions followed. First, participants still projected more positive trends after a positive (vs. negative) price change, even if told that the change was random: People are "fooled by randomness" (Taleb, 2001; Tversky & Kahneman, 1971), even when randomness is noted explicitly. Second, the no-explanation condition was always more extreme than the noise condition, but less extreme than the internal-explanation condition: People consider unexplained price changes to contain some signal but not as much as explained changes. Third, this effect was asymmetric. For positive changes, the no-explanation was closer to the internal-explanation condition, whereas for negative changes, the no-explanation was closer to the noise condition. Thus, unexplained price changes – accounting for most volatility – are treated more like signal when positive and like noise when negative. This could lead to bubble dynamics, wherein positive price trends build on themselves but the corresponding force in the negative direction is weaker.

#### 7.1.2. Valence structure: approach and avoidance

Events in narratives often have positive or negative valences, motivating approach or avoidance behavior. To test the effect of valence structure on forecasts, participants read about fictitious companies which experienced either good (an increase in earnings) or bad news (a decrease in new oil discoveries), announced prior to the most recent stock price quotation (Johnson & Tuckett, 2022). According to financial theory, there is no further effect on the stock price once the information is publicly revealed and priced in. However, if people use narratives to organize their beliefs, they should continue to rely on this valence information to predict future value.

They did. Without news, participants thought that a stock would increase in value by +4.3% in the following two weeks. These predictions were much more extreme when news was available (+10.1% and −5.9% for good vs. bad news). This implies that people believe that markets profoundly underreact to news, leading to price momentum (prices trending in a particular direction). Could participants have intuited the finding that financial markets *do* experience modest price momentum in the short- to medium-term after news announcements, which then reverts back toward the baseline trend (Shefrin, 2002)? If so, they should predict smaller gaps between positive and negative news at longer intervals. In fact, the predicted gap is *larger* at a one-year interval (+16.1% vs. −5.9%). Price expectations seem to follow valenced stories about the companies' underlying causal propensities – with companies categorized as "good" versus "bad" – rather than economic intuition.

#### 7.1.3. Temporal structure: asymmetries between past and future

Narratives also contain temporal structure, which implies a boundary condition on the effect of valence: When information is related to the future (vs. the past), we should see more of an effect of its valence on predictions. Indeed, consistent with CNT but not standard financial theory, news about the future (e.g., a revision to next quarter's projected earnings; +17.5% predicted one-year change) stimulated more extreme predictions compared to news about the past (e.g., a revision to last quarter's actual earnings; +14.7%). Thus, not only the valence but also the temporal orientation of news affects forecasts, in line with narrative representations (Johnson & Tuckett, 2022).

Moreover, the valence- and time-based predictions described in Sections 7.1.2 and 7.1.3 manifested in emotions – more positive forecasts led to more approach emotions, more negative forecasts to more avoidance emotions – which in turn motivated investment decisions. This confirms a basic principle of CNT: People use narratives to imagine the future, react affectively to that future, and choose in line with their affect (Section 8.1).

#### 7.1.4. Analogical structure: pattern detection and extrapolation

Analogies allow us to impose structure on problems by using our knowledge of one thing to understand another. Although some analogies compare radically different things (e.g., an atom is like the solar system), most analogies are much more prosaic: This dishwasher is like that dishwasher, this dog behaves like other dogs, this company's future resembles that other company's.

Our minds seem to hold multiple, conflicting analogies which can impose structure on time-series price data. On the one hand, it seems plausible that a series of prices trending in one direction should continue that trend – *momentum*. Many familiar variables have this property – successful people tend to become even more successful; objects in motion tend to stay in motion (as in the analogy of "momentum" itself). At the same time, it seems equally plausible that if prices have trended one way, it's only a matter of time before the trend reverses – *mean reversion*. Mean reversion too is common among familiar variables – extreme weather regresses toward the mean; objects thrown into the air come down eventually. The extent of momentum and mean reversion in real prices has been a matter of great debate in behavioral finance. Given that people can harness sophisticated intuitions for pattern detection and extrapolation (Bott & Heit, 2004; DeLosh, Busemeyer, & McDaniel, 1997), how might we

https://doi.org/10.1017/S0140525X22001157 Published online by Cambridge University Press

resolve these conflicting analogical intuitions when predicting future prices?

We anticipated that these analogies are triggered by different evidence, with these analogical frames imposed on time-series data to best explain it and project that explanation into the future. This signature bias would contrast with standard financial theory, which says that only the current price and its variance (risk) are relevant to future prices, as well as with existing behavioral models which assume that people are linear trend extrapolators (Barberis, Greenwood, Jin, & Shleifer, 2015).

In our studies, participants encountered prices series in one of three patterns (Johnson, Matiashvili, & Tuckett, 2019b). In the *linear* condition, the prices had been trending in either a positive or negative direction for five periods. Here, participants linearly extrapolate the trend, consistent with past work (Cutler, Poterba, & Summers, 1991; De Bondt, 1993; Hommes, 2011; Jegadeesh & Titman, 1993). But in two other conditions, we find strikingly different results. In the *reversion* condition, prices had previously experienced a reversion during the past 5 periods – they had the same starting price, ending price, and mean as the linear condition, but experienced higher variance in the intervening periods. Participants had a greatly dampened tendency to project this trend linearly, with many believing that prices would reverse again. In the *stable* condition, prices had previously hovered around one price level before experiencing a sudden increase or decrease to the current price. This pattern too led many to predict reversion, toward the previously stable price.

These results suggest that people draw on analogies – such as momentum and mean reversion in other data series – to generate narratives to account for past price trends, projecting these narratives to forecast the future. We find similar pattern-based expectations for many other consumer and investment prices, and real stock prices in an incentive-compatible task. Beyond their theoretical implications, these results may be economically significant, as price expectations play key roles in asset pricing (Hommes, Sonnemans, Tuinstra, & van de Velden, 2005) and inflation (Carlson & Parkin, 1975).

### 7.2. Imagined futures are simulated one at a time

We have focused so far on how narratives circumvent limitations on probabilistic thinking. Yet narrative thinking has limits of its own: Once we have adopted a particular narrative, we are often blind to alternative possibilities. We simulate narratives one at a time.

For example, different government policies lead to different predictions about market prices. If the central bank raises interest rates, this is likely to depress market prices. But central bankers often speak opaquely. If an investor assigns a 75% chance to the story that the banker plans to raise rates but a 25% chance to the story that the banker plans to leave them alone, does she account for both possibilities or rely on just one? A Bayesian would calculate the likely effect on markets if each story is true, taking a weighted average when estimating future prices. But an investor who "digitizes" and treats these stories as either certainly true or false would "round up" the 75% chance to 100% and "round down" the 25% chance to 0%. Only the dominant narrative resulting from explanatory reasoning would be retained for downstream computations such as forecasting.

We found that investors are not Bayesians, instead digitizing (Johnson & Hill, 2017). In one study, participants were given information leading them to think that one government policy was likelier than another (in one variation, they were even given these probabilities directly). Comparing conditions identical except the effect of the more-likely (75% chance) policy, there was a large difference in predictions. But comparing conditions identical except the effect of the less-likely (25% chance) policy, predictions do not differ at all. Investors take account of the implications of more-likely narratives, but ignore entirely the implications of less-likely narratives: They adopt a single narrative as true, treating it as certain rather than probable.

Digitization is a broad feature of cognition; similar effects have been found in causal reasoning (Doherty, Chadwick, Caravan, Barr, & Mynatt, 1996; Fernbach, Darlow, & Sloman, 2010; Johnson, Merchant, & Keil, 2020) and categorization (Lagnado & Shanks, 2003; Murphy & Ross, 1994). Yet it has boundary conditions: People do reason across multiple hypotheses in cases where one of the hypotheses invokes a moral violation (Johnson, Murphy, Rodrigues, & Keil, 2019) or danger (Zhu & Murphy, 2013), and expertise in a domain may promote multiple-hypothesis use (Hayes & Chen, 2008).

Despite boundary conditions, simulations produce, by default, a single imagined future. Electrons may exist as probability clouds rather than in one definite state. But stories resist Heisenberg's principle – they take only one state at a time.

## 8. Affective evaluation

We have seen how narratives solve the mediation problem: They summarize available data (about the past) in a format used to predict what will happen (in the future) given a particular choice. But choices must somehow combine these visions of the future with our values and goals – the combination problem. Since emotions function to coordinate goals, plans, and actions (Damasio, 1994; Elliot, 2006; Fishbach & Dhar, 2007; Ford, 1992; Lewin, 1935; Oatley & Johnson-Laird, 1987; Rolls, 2014), *affective evaluation* is tasked with solving the combination problem.

Let us first consider how existing theories of emotion address simpler choices. Suppose someone cuts into the supermarket queue and you must decide whether to assert your rightful place. According to the *appraisal-tendency framework* (Lerner & Keltner, 2000), the decision-maker evaluates this situation along several dimensions – including certainty, pleasantness, controllability, and others' responsibility – which jointly determine which emotion is felt (Smith & Ellsworth, 1985). In the queue-cutting case, one might perceive the event as unpleasant and the queue-cutter as responsible, but the situation as certain and under control – leading to anger; or instead, one might be less sure that the queue-cutter acted deliberately but perceive the situation as less controllable and certain because the queue-cutter appears big and mean – leading to fear. These emotions, in turn, motivate different actions (Frijda, 1988); anger is an approach emotion motivating aggression, whereas fear is an avoidance emotion motivating withdrawal. Finally, once these emotions are active, they shift attention to relevant dimensions for subsequent events; for example, fear leads one to perceive subsequent events as relatively uncontrollable compared to anger (Lerner & Keltner, 2001).

CNT adds three modifications to account for challenges in complex, future-oriented decision-making. First, emotions are felt not only in response to actual events but to *imagined futures* generated from narratives, motivating us to approach or avoid associated choices. Second, appraisals of those futures can rely either on a default set of dimensions or on ad hoc evaluations relative to specific goals. Finally, since decisions must often be sustained over time, feelings of conviction in a narrative permit committed action in the face of uncertainty.

### 8.1. Affective evaluations of imagined futures motivate choices

We feel emotions in response not only to the present situation, but to situations we imagine (Loewenstein, Weber, Hsee, & Welch, 2001; Richard, van der Pligt, & de Vries, 1996). This is why we experience emotions when understanding literature (Mar, Oatley, Dikic, & Mullin, 2010) or empathizing with others (Mitchell, Banaji, & Macrae, 2005). CNT accords a central role to the emotional reactions we experience in response to imagined futures generated from narratives. These emotional reactions *within* a narrative, by motivating approach and avoidance behaviors, drive action in the real world.

Some emotions are inherently future-oriented: If one feels excited (or anxious) about a potential future, one acts to approach (or avoid) that future. But even past-oriented emotions, such as regret (Loomes & Sugden, 1982), can influence our choices through simulations of how we would feel. For example, people anticipate more regret over not playing in postcode lotteries (where non-players can learn if they would have won) versus traditional lotteries, which motivate participation (Zeelenberg & Pieters, 2004). Many other anticipated emotions are known to guide choices, including guilt, sadness, anger (Baron, 1992); pleasure (Wilson & Gilbert, 2005); dread and savoring (Dawson & Johnson, 2021; Loewenstein, 1987); and envy (Loewenstein, Thompson, & Bazerman, 1989). Emotions mediate between our predictions of the future and decisions to approach or avoid that future, coloring narratives with emotion (Beach, 1998; Bruner, 1986).

### 8.2. Imagined futures can be appraised on default or ad hoc dimensions

The fuzzy evaluation problem (Section 2.2.2) results from the challenges of summarizing incommensurable attributes as a single utility, especially when our values may change over time. CNT proposes two computational simplification strategies that the affective system can use to address this problem – a default, bottom-up strategy and an ad hoc, top-down strategy.

This is analogous to the distinction between natural categories and ad hoc categories. Natural categories – such as BIRD, TABLE, and MOUNTAIN – roughly capture regularities in the external world (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976); by default, objects are classified bottom-up, automatically and effortlessly, into natural categories (Greene & Fei-Fei, 2014). In contrast, ad hoc categories (Barsalou, 1983) – such as THINGS TO SELL IN A GARAGE SALE and WAYS TO ESCAPE THE MAFIA – are constructed on the fly to achieve specific goals, using effortful, top-down processes. Whereas bottom-up classification into natural categories proceeds by default and relies on predetermined dimensions, top-down classification into ad hoc categories requires effort and relies on spontaneously determined, goal-derived dimensions.

Analogously, in line with the appraisal tendency framework, one strategy for evaluating imagined futures relies on a default set of dimensions mirroring those for evaluating actually present situations (e.g., controllability, certainty, pleasantness), which determine which emotion is felt. That emotion, in turn, motivates action. Because these dimensions are thought to be evaluated automatically with minimal effort (Lazarus, 1991), this default appraisal strategy is an appealing solution to the fuzzy evaluation problem. Specific emotions are felt in response to qualitative appraisals of predetermined dimensions. Because the dimensions are predetermined, the computational problem of identifying dimensions is avoided; because the appraisals are qualitative, the need to trade off these dimensions is circumvented. Moreover, although particular preferences may well change over time, our basic emotional architecture does not. Thus, the problems of incommensurable attributes and non-stationary values are averted.

However, these default dimensions often do not suffice when we have specific goals, leading to a second, ad hoc strategy based on the decision-maker's goal hierarchy. A decision-maker's attention will be deployed according to the active goal(s) at the time of decision-making (Van Osselaer & Janiszewski, 2012). Narratives can be used to generate imagined futures, on an ad hoc basis, that are evaluable with respect to these goals. The compatibility of those imagined futures with those goals produces approach and avoidance emotions that motivate action (Elliot, 2006; Oatley & Johnson-Laird, 1987).

This ad hoc route depends on two claims. First, we assume that although we may have many goals, a small subset are typically active at once, because goals are triggered context-dependently (Panksepp, 1998; Tuckett, in press) and organized hierarchically. For example, when basic physiological needs are not met, these are likely to supersede less immediately essential needs such as social belonging (Lavoie, 1994; Maslow, 1943). This is not only helpful for survival, but also a crucial computational simplification: Goal hierarchies allow us to evaluate imagined futures over a much smaller number of dimensions. This is how ad hoc appraisals, like default appraisals, help to resolve the fuzzy evaluation puzzle. This insight also casts new light on multi-attribute choice strategies (Payne, Bettman, & Johnson, 1988). A couple facing divorce faces a dizzying array of potential attributes. Yet, for many such couples, their children's well-being is paramount. If this dimension does not prove decisive, they may move on to other concerns, such as their financial well-being or sexual satisfaction. The hierarchical organization of goals can explain why particular situations call for particular decision rules, such as lexicographic rules (using a single attribute) or elimination-by-aspects (eliminating options beneath a minimum criterion on a key attribute, then iteratively considering other dimensions; Tversky, 1972).

Second, we assume that we can generate imagined futures containing details relevant to evaluating the required dimensions. For example, our married couple might first imagine their future with respect to their child's well-being, then elaborate this image to consider their romantic prospects, then their finances. Although these different imagined aspects may well emanate from a shared narrative of the couple's married life, it is unlikely that this complete conception of their future would emerge fully formed, but instead must be simulated one piece at a time. Although we do not know of any research directly examining this proposed ability in the context of narratives, evidence about other domains suggests that this is possible; for example, people can fluidly reclassify objects dependent on goals (Barsalou, 1983) and can manipulate their mental images dependent on queries (Kosslyn, 1975).

### 8.3. Emotions are used to manage decisions extended over time

Hamlet's uncertainty paralyzed him for three acts; by the fifth act, it was too late. Hamlet learned the hard way that strong, conflicting arguments produce ambivalence that can stop action in its

https://doi.org/10.1017/S0140525X22001157 Published online by Cambridge University Press

tracks (Armitage & Conner, 2000; Festinger, 1962; Rosner et al., 2022; Rucker, Tormala, Pety, & Briñol, 2014; Smelser, 1998). Many a couple and many an investor have talked themselves in circles while romantic and financial opportunities slipped away; committing to one distinct course of action often yields better fortunes, even if one can never be certain a choice is truly "right." Moreover, high-stakes decisions often are extended through time, requiring commitment. Failure in this respect leads many novice investors to overtrade and defray their gains through transaction costs (Barber & Odean, 2000). Conviction bears dividends.

Yet conviction also bears risks. Blindly following a plan, while ignoring new information, is equally a recipe for disaster as Hamletian paralysis. Complex problems such as the COVID pandemic, Putin's Ukraine invasion, or climate change require different approaches as events or our knowledge evolve. Emotions are instrumental in the inter-related processes of *conviction management* – *gaining conviction* (acting in the face of ambivalence), *maintaining conviction* (committing to a sustained course of action), and *moderating conviction* (taking account of new evidence and potentially changing course as the situation changes). To manage conviction is to manage our emotional attachments to a person, object, or course of action. A lack of emotional attachment against the temporary vicissitudes of fortune yields indecision, yet an inability to reappraise a truly changing situation can yield calamity.

Cognition and affect are intertwined in generating conviction. Conviction-generating narratives integrate information about the past and expectations about the future to emotionally support a course of action. Experiments have probed this process. In a purely cognitive model, confidence in a decision is proportional to the strength of the arguments in its favor; in a purely affective model, confidence in a decision is proportional to its propensity to trigger approach emotions. Instead, cognition and emotion work together (Bilovich, Johnson, & Tuckett, 2020). When situations trigger approach emotions, investors find favorable arguments more relevant, with the converse for avoidance emotions. Perceived relevance in turn influences investors' choices. Although emotions profoundly affect choice, they do so by influencing intermediate cognitive processing.

This interplay between cognition and affect is also illustrated by the conviction-generating strategies cited by the investment managers in Tuckett's (2011) interview study (Chong & Tuckett, 2015). Numerous respondents (90%) referred to one or more "attractor" narratives, producing excitement over an investment due to an exceptional opportunity for gain. Attractor narratives typically cite either the investment's intrinsic properties (e.g., exceptional products) or the investor's own special skills (e.g., exceptional insight). A similar proportion of respondents (88%) referred to one or more "doubt-repelling" narratives, reducing anxiety over an investment. Doubt-repelling narratives typically raise and then counter a potential concern, placing bounds on either uncertainty (e.g., competent managers) or downside surprise (e.g., solid fundamentals).

Uncertainty can undermine conviction, but so can excessive certainty in dynamic situations. At a single time-point, people are more confident in investments described as having a specific, predictable return (8%) rather than falling within a range (3–13%) (Batteux, Bilovich, Johnson, & Tuckett, in press; Du & Budescu, 2005). However, building conviction by masking uncertainty is not sustainable: Once point forecasts are shown to be unreliable – an inevitable event under uncertainty – trust is reduced in the forecaster (Batteux, Bilovich, Johnson, & Tuckett, 2021a). This has implications for risk communication: When uncertainty is communicated in vaccine announcements, trust in vaccines is buffered against subsequent negative outcomes (Batteux, Bilovich, Johnson, & Tuckett, 2021b).

Conviction is not good or bad in itself. It is needed to overcome ambivalence and sustain commitment, but is only adaptive if it does not preclude learning. When emotions are regulated well (Gross, 1998), conviction buffers against the vicissitudes of the unfolding situation, but can be moderated. In such an *integrated state*, we can be sensitive to new information and adjust decisions in an orderly way; one adopts a particular narrative but acknowledges the possibility of error and stays attuned to evidence for competing narratives. In contrast, in a *divided state*, ambivalence is hidden by attentional neglect of information inconsistent with the preferred narrative. Whereas new information in integrated states can trigger curiosity and evidence integration, incongruent information is rejected in a divided state (Tuckett, 2011; Tuckett & Taffler, 2008). This is why ambivalence has been linked both to maladaptive responses, such as confirmation bias and behavioral paralysis, and adaptive responses, such as broader attention and willingness to consider multiple perspectives (Rothman, Pratt, Rees, & Vogus, 2017). Decision-makers who experience balanced emotions are likelier to rely on "wise reasoning" strategies, such as epistemic humility and integrating diverse perspectives (Grossmann, Oakes, & Santos, 2019). Integrated conviction management is adaptive because it permits decision-makers to accept a narrative as provisionally true and act accordingly – a crucial characteristic when there is a cost to changing course – while accumulating evidence in the background, changing course when clearly merited.

Integrated conviction management is closely linked with one way that feedback loops help decisions to become adaptive over time. Although under radical uncertainty we often have no choice but to make some decision without a clear sense of whether it is the best option, we can accumulate evidence about what does and does not work. Thus, *acting on* one narrative can yield information that is then used to *reappraise* that same narrative, potentially leading to a shift in narrative and decision in an iterative manner. Extended decisions can often be treated as a series of experiments, providing information about what does and does not work (Fenton O'Creevy & Tuckett, 2022).

## 9. Communication

Many everyday decisions are inseparable from their social context. The *communication* processes through which narratives or narrative fragments are transmitted across minds are crucial for understanding decision-making at macro scales. Subjecting narratives to cultural evolution has allowed narratives to adapt over time to generate reasonably high-quality decisions in the absence of calculable probabilities and utilities.

### 9.1. Shared narratives facilitate social coordination

Decisions are socially embedded in part through their shared consequences – when decisions are taken collectively, as in political decision-making, or when one individual's decision affects others in their social group. Socially coordinated decisions can generate more value than the sum of their individual components (Chwe, 2001). However, coordination is challenging, due to both divergent interests and divergent information. Shared narratives help to coordinate both interests and information.

https://doi.org/10.1017/S0140525X22001157 Published online by Cambridge University Press

Reputation-tracking is key to aligning individual incentives with collective interest (Rand & Nowak, 2013; Tennie, Frith, & Frith, 2010). People are motivated to evaluate others' reputation based on their actions, even actions only affecting third parties. For instance, people evaluate others' moral character based on prosocial actions such as donation (Glazer & Konrad, 1996; Johnson, 2020), volunteering (Johnson & Park, 2021), and eco-friendly actions (Griskevicius, Tybur, & van den Bergh, 2010), generating incentives for apparent altruism. Conversely, bad reputations are costly: People sacrifice resources to punish free-riders (Jordan, Hoffman, Bloom, & Rand, 2016). Because we are aware that others are tracking our reputations, we are motivated to take actions bringing reputational benefits and avoid actions bringing reputational harm.

An important means for reputation management is how we justify choices to others (Lerner & Tetlock, 1999; Mercier & Sperber, 2017). Narratives often play this justificatory role, maintaining reputation in the face of disagreement and coordinating group activity when other stakeholders must adopt the same decision. For instance, stories shared by the Agta, hunter–gatherers in the Philippines, express messages promoting cooperation, allowing skilled storytellers to achieve greater cooperation (Smith et al., 2017). Closer to home, scientists often debate what "story" they will sell to readers and reviewers. Why are narratives so effective for reputation maintenance?

First, narratives contain causal structure that can generate *reasons* justifying a position. For example, when deciding which parent should be awarded custody of a child, people favor the parent with both more extreme positive (above-average income) and negative attributes (work-related travel) over one with more neutral attributes (typical income and working hours), since the former gives more positive reasons favoring custody. But when asked instead who should be *denied* custody, people again choose the more extreme parent because there are also more *negative* reasons *against* custody (Shafir, Simonson, & Tversky, 1993). More extreme attributes support more causally potent explanations, generating both supporting and opposing narratives.

Second, narratives can not only justify decisions after the fact but to *persuade* others to adopt our perspective (Krause & Rucker, 2020). Arguments communicated with a narrative are often more persuasive than those communicated with facts alone (Chang, 2008; De Wit, Das, & Vet, 2008; Shen, Ahern, & Baker, 2014), in part because narratives are readily understood (Section 9.2). Narratives induce emotional engagement, mental imagery, and attention, creating "narrative transportation" that can lead reasoners to believe elements of the story (Adaval & Wyer, 1998; Escalas, 2004; Green & Brock, 2000; Hamby, Brinberg, & Daniloski, 2017; Van Laer, de Ruyter, Visconti, & Wetzels, 2014). Moreover, the broader narratives espoused by a communicator, such as moral and political worldviews, can lend additional credence to their claims (Johnson, Rodrigues, & Tuckett, 2021; Marks, Copland, Loh, Sunstein, & Sharto, 2019). Persuasion is crucial in coordination because it allows a group to have the *same* narrative in their heads, making narratives a part of our collective or transactive memory (Boyd, 2009; Chwe, 2001; Hirst, Yamashiro, & Coman, 2018; Wegner, 1987) and providing a shared plan for coordinated action.

### 9.2. Shared narratives shape social learning and evolve

Decision-making is also socially embedded through the informational environment. Human knowledge arises largely through our cumulative cultural heritage (Boyd, Richerson, & Henrich, 2011; Henrich, 2018). Indeed, because it is so often the ability to access external knowledge when needed that is crucial for decision-making rather than our internal knowledge itself, we often confuse knowledge inside and outside our heads (Sloman & Fernbach, 2017).

Communication of narratives is a crucial way we learn beyond immediate experience (Boyd, 2018), with some even suggesting that the adaptive advantage of sharing narratives is the main reason that language evolved (Donald, 1991). However, we do not suggest that narratives in the full form described in Section 5 migrate wholesale from one mind to another. The knowledge that is transferred from one mind and stored in another is relatively skeletal, or even a placeholder (Rozenblit & Keil, 2002) as when we store the *source* of a piece of information rather than the information itself (Sloman & Fernbach, 2017; Sparrow, Liu, & Wegner, 2011). Instead of full narrative representations, we assume instead that primitive elements – narrative fragments such as basic causal schemas, memorable analogies, and emotional color – are the key narrative elements that are shared and shape social learning. These narrative fragments, communicated consistently within a social group, give rise to a set of elements that are common among the narratives represented within those group members' individual minds – what we are calling a *shared narrative*.

Several approaches seek to model how ideas spread and evolve (Boyd & Richerson, 1985; Dawkins, 1976; Sperber, 1996). A common insight is that ideas spread when they pass through two sets of cognitive filters: Constraints on encoding (attention, memory, and trust) and constraints on communication (motivation and ability to share). But for a culturally transmitted idea to not only spread but be *acted on*, that idea must pass through a third filter – constraints on action. The idea must be perceived as actionable and produce motivation to act. Narratives can pass through all three filters – encoding, communication, and action – and therefore are likely to socially propagate.

First, narratives are easy to remember. People often represent information as scripts, or stereotyped sequences of events (Schank & Abelson, 1977). Because people so naturally represent information using causal–temporal structure, people are far better at remembering information organized as stories (Bartlett, 1932; Kintsch, Mandel, & Kozminsky, 1977; Thorndyke, 1977). Humans' remarkable ability to remember information encoded as stories is perhaps most impressively attested by oral traditions, such as the transmission of the Greek epics and Hindu and Buddhist historical texts over centuries purely through word-of-mouth (Rubin, 1995).

Second, people like to talk about narratives, making them susceptible to spreading through word-of-mouth (Berger, 2013). This may be because narratives are well-suited to balancing novelty against comprehensibility (Berlyne, 1960; Silvia, 2008), with the most contagious narratives including a small number of novel concepts against a larger background of familiar concepts (Norenzayan, Atran, Faulkner, & Schaller, 2006). Narratives can convey new information in digestible form because they match our default causal–temporal representations of events (Schank & Abelson, 1977) and focus on the behavior of human actors, commandeering our natural tendency toward gossip (Dunbar, 1996). Moreover, as discussed above, narratives are highly persuasive, and therefore commonly used when trying to convince others.

Finally, narratives lend themselves to action. Because narratives have a causal–temporal organization, and are often organized around the actions of individuals (Mandler & Johnson, 1977), they provide a ready template for intervening on the world.

https://doi.org/10.1017/S0140525X22001157 Published online by Cambridge University Press

Indeed, causal knowledge is crucial precisely because it can be used to bring about desired outcomes (Woodward, 2003). Culturally acquired knowledge of physical causation is embedded in physical tools, which can manipulate the physical world. Analogously, culturally acquired knowledge of social causation is embedded in narratives, which serve as templates for manipulating the social world. We suggest that narratives which lead to *effective* actions are particularly likely to survive this filter, just as effective institutions are likelier to survive social evolution (Hayek, 1958). Therefore, narratives that have survived this crucible of cultural evolution potentially lead to adaptive decision-making even under radical uncertainty.

Because narratives live and die by cultural evolution (Boyd & Richerson, 1985; Henrich, 2018), their propagation depends on their social, economic, and physical environment. Narratives surrounding masks at the start of the COVID-19 differed profoundly across countries (Hahn & Bhadun, 2021), due both to social norms and differing experience with infectious disease. Henrich's (2020) account of how the West became prosperous suggests that narratives originated by the Catholic Church altered norms around cousin marriage, generating new family structures and patterns of cooperation that led to markets and science.

### 9.3. Shared narratives propagate through social networks

If shared narratives facilitate coordination and learning, then as they shift over time and propagate through social networks, they should be tied to large-scale outcomes. Since economic actors' decisions are driven by narratives (Section 5.2.2), changes in the emotional content of socially available narratives may shift attitudes toward risk. If so, then measures of approach and avoidance emotions in economic narratives should predict the direction of economic aggregates – output, employment, GDP growth – that depend on investment. Nyman, Kapadia, and Tuckett (2021) studied this claim using text-mining techniques on internal Bank of England commentary (2000–2010), broker research reports (2010–2013), and Reuters news articles (1996–2014).

First, relative sentiment is a leading indicator of economic volatility and consumer sentiment. This reflects the idea that a preponderance of approach over avoidance emotions is needed to produce conviction to invest (Keynes, 1936). For each document, the proportion of words signaling approach (e.g., "excited," "ideal") and avoidance words ("threatening," "eroding") was calculated, with the difference between these indices constituting *relative sentiment*. Shocks to relative sentiment in the UK had negative effects on industrial production, employment, and the stock market, with these impacts lasting for nearly 20 months (Nyman et al., 2021). Tuckett and Nyman (2018) also found that relative sentiment also predicted changes in investment and employment in the UK, US, and Canada more than 12 months out (Tuckett, 2017). For example, a plot of relative sentiment against major events in the lead-up to the global financial crisis shows a precipitous decline in relative sentiment in the year leading up to the failure of Bear Sterns in March 2008. A similar analysis using 1920s data from the *Wall Street Journal* found that sentiment shocks, beyond economic fundamentals, impacted production and stock values leading up to the Great Depression (Kabiri, James, Landon-Lane, Tuckett, & Nyman, 2023); sentiment likewise appears to account for the slow recovery from the 2008 recession (Carlin & Soskice, 2018).

Second, excessive homogeneity around narratives portends trouble. Nyman et al. (2021) used topic modeling to assign each article in the text database to a particular narrative and compute the degree of narrative topic consensus at each time point. Prior to the crisis, homogeneity increased around narratives high in approach emotions (excitement) and lacking avoidance emotions (anxiety), which could have been a potential warning sign of impending financial system distress (Nyman et al., 2021). This supports the idea of groupfeel or emotional conformity as a driver of booms and busts and the notion that integrated emotional states that manage ambivalence are better-suited to stable decision-making, compared to divided states that ignore discordant information (Tuckett & Taffler, 2008).

Macroeconomic crises are necessarily rare and atypical, so no method will reveal definitive answers about their causes. But these techniques can also be used at a more micro-scale. For example, Tuckett, Smith, and Nyman (2014) studied relative sentiment in news articles about Fannie Mae. From 2005 to mid-2007, sentiment became increasingly exuberant, along with Fannie Mae's share price, and unmoored from economic realities reflected in the Case-Shiller Housing Price index. Such states can result from the fetishization of some "phantastic object" (Tuckett & Taffler, 2008) – in this case, mortgage securitization. A similar analysis of Enron's internal emails in 2000–2002 revealed comparable emotional–narrative dynamics of build-up and collapse surrounding the deregulation of the California energy market and Enron's impending (ill-fated) entry into broadband. Overall, the confluence of macro- and micro-level analyses converges to suggest that emotional–narrative sentiment spreads through social networks and may causally influence economic outcomes.

## 10. Conclusion

So often, decisions in economics textbooks and psychology laboratories alike are divorced from the need for sense-making and imagination, overtly quantified in their consequences and probabilities, taken at a single time-point, and stripped of social context. This reductionist tradition has yielded massive progress. But progress in understanding everyday decision-making also requires us to put back those elements that have been stripped away. CNT is our answer to this need.

To summarize CNT standing on one foot: *We impose narrative structure on information to explain the past, imagine the future, appraise that future, take sustained action, and coordinate actions socially.* The mediation problem – a mental representation that can mediate between the external world and our choices – is solved by narratives; the combination problem – a mental process that can drive action by combining beliefs and values – is solved by emotion. Decisions can be reasonably adaptive even without well-specified probabilities and utilities because individual narratives are influenced by culturally evolved shared narratives and by feedback from our own actions.

### 10.1. Meta-theoretical considerations

Scientific theories, like narratives, are evaluated partly on aesthetic grounds. In this spirit, we discuss two potential "meta-theoretical" objections to CNT.

First, some may view CNT as too *grandiose*. Philosophers have mostly given up on generating grand unified theories, and similar efforts such as behaviorism have fared little better in psychology. A skeptical reader may view CNT as a kaleidoscope of ideas – encompassing narratives, explanation, causation, analogy, forecasting, emotion, motivation, cultural evolution, and more – biting off too much theoretical meat to properly chew, much less digest.

Second, some may view CNT as too *skeletal*. We do not provide our own account of explanation or emotion or cultural evolution, but rather focus on how these processes fit together. Even the notion of narratives – the theoretical centerpiece of our account, as the mental substrate binding these processes – can be elusive, as it coordinates lower-level representations of causation, analogy, time, and valence. Perhaps we have not bitten off theoretical meat at all, but merely bones.

We are sympathetic to these concerns. Yet we believe that in this case, we have encountered a grand problem that *requires* a strong skeleton. CNT is not grand in the sense that it attempts to explain all of cognition; rather, it is attempting to explain decision-making under radical uncertainty – an important problem that has largely resisted theoretical progress. It is a grand problem precisely *because*, we contend, many parts of our mind cooperate to make such decisions. Approaches that ignore any component piece will lack the theoretical machinery required to understand the problem. If CNT is grand, it is not by choice but by necessity.

Thus, CNT is *not* a theory of explanation, analogy, causation, or emotion, but a theory of *decision-making under radical uncertainty*. We focus less on the details of these component processes because it is how the processes *interact* that is central to how we produce the conviction to act under uncertainty. We do not necessarily take a side where there is theoretical disagreement about one of these processes; rather, we specify how these processes relate *to each other*. CNT is skeletal because the skeleton *is* the theory; the meat, though delicious, belongs to other theories.

### 10.2. Contributions

We are not the first to highlight the importance of narratives to decision-making, nor (we hope!) the last. But we have aimed to provide an integrative framework that allows insights from several disciplines to be combined, contributing to the ongoing conversation in four main ways.

First, by explaining in detail how narrative decision-making works. We provide a representational framework that captures key information used in real-world examples of narrative-based decisions, and explain how these representations sustain processes of explanation, simulation, and affective evaluation, which jointly motivate action.

Second, by highlighting that narratives address important puzzles about everyday decision-making. Many ordinary decisions are plagued by radical uncertainty, fuzzy evaluation, and the need for sustained commitment; they involve sense-making and imagination; they are inextricably linked to social context. Such decisions – where optimality is ill-defined – resist dichotomization into "rational" versus "irrational." Narratives not only help to solve these problems, but often do so adaptively. Feedback loops at both the individual level (managing conviction for decisions sustained over time) and the collective level (the cultural evolution of narratives) contribute to adaptive choice.

Third, by identifying how processes ordinarily studied in isolation work together. Recent advances in explanatory reasoning highlight the role of heuristics and affect in explanation under radical uncertainty – advances unknown at the time of Pennington and Hastie's (1986) seminal work. Causal and analogical processing have been studied extensively, with excellent cognitive models of each; yet these models have been integrated surprisingly little with one another or with decision-making models. The pivotal role of affect in solving the fuzzy evaluation problem has received less than its deserved attention in the decision-making literature, as has the role of cultural evolution in socially propagating narratives that can then guide individual choices. We hope to put these areas in dialogue.

Finally, by providing a common vocabulary – including our ideas around the structure of narrative representations, the information flow among narrative processes, and the set of problems to be addressed by a theory of everyday choice.

In keeping with this final point, we emphasize that a common vocabulary is needed precisely because CNT will not be the final word on this topic. Indeed, some of our proposals remain tentative even for us. First, although we believe that causal, temporal, analogical, and valence structure are the key lower-level information included in narratives, there may be other forms of information we have not considered. Second, relatively little is known about how these kinds of information are coordinated; thus, our proposals around narrative coherence rules must remain tentative and likely incomplete. Third, although much has recently been learned about explanatory heuristics, we have not provided an exhaustive list of these heuristics but merely given a few examples to illustrate how they work; more will surely be discovered. Fourth, we know relatively little about what specific features of narratives lead them to be more or less socially contagious and which of those features promote adaptive decisions; memes are selfish, as Dawkins (1976) noted, and not all features of catchy narratives are likely to be adaptive. Yet we would add that the preoccupation of much decision-making research with optimality – whether in assumption or subversion – might profitably yield some ground to the more basic question of how, under radical uncertainty and fuzzy evaluation, we gain conviction to act at all.

We are excited by the prospect that CNT might provide a fruitful platform for collaboration between researchers across the decision sciences – a rallying cry for all who aim to understand social, psychological, and economic aspects of decision-making in the real world.

**Acknowledgements.** We thank Liz Allison, Eleonore Batteux, Tim Besley, Gordon Brown, Christopher Dawson, Arman Eshraghi, Mark Fenton O'Creevy, Peter Fonagy, Aikaterini Fotopoulou, Karl Friston, Gerd Gigerenzer, Reid Hastie, Douglas Holmes, John Kay, Mervyn King, Dave Lagnado, Rickard Nyman, David Shanks, George Soros, Lukasz Walasek, and the members of UCL's Centre for the Study of Decision-Making Uncertainty, for useful comments, encouragement, and discussion.

**Financial support.** This research was supported by grants to DT and the UCL Centre for the Study of Decision-Making Uncertainty from the Institute of New Economic Thinking (IN011-00025, IN13-00051, and INO16-00011), the Eric Simenhauer Foundation of the Institute of Psychoanalysis (London), UKRI (EPSRC grant reference EP/P016847/1), and the ESRC-NIESR Rebuilding Macroeconomics network.

**Competing Interest.** None.

## Note

1. Excerpts can be found in Tuckett (2012). The full interview, along with the three others selected at random from this larger group, is available at https://www.macmillanihe.com/companion/Tuckett-Minding-The-Markets/study-resources/. More detailed analysis of all the decisions reported by the entire sample, also supported by randomly drawn examples from the coded interview data, has been reported elsewhere (Tuckett, 2011).

https://doi.org/10.1017/S0140525X22001157 Published online by Cambridge University Press

## References


Abbott, V., Black, J. B., & Smith, E. E. (1985). The representation of scripts in memory. *Journal of Memory and Language*, *24*, 179–199.

Abolafia, M. Y. (2020). *Stewards of the market: How the Federal Reserve made sense of the financial crisis*. Harvard University Press.

Adaval, R., & Wyer, R. S. (1998). The role of narratives in consumer information processing. *Journal of Consumer Psychology*, *7*, 207–245.

Ajzen, I. (1977). Intuitive theories of events and the effects of base-rate information on prediction. *Journal of Personality and Social Psychology*, *35*, 303–314.

Akerlof, G. A., & Shiller, R. J. (2009). *Animal spirits: How human psychology drives the economy, and why it matters for global capitalism*. Princeton University Press.

Akerlof, G. A., & Snower, D. J. (2016). Bread and bullets. *Journal of Economic Behavior & Organization*, *126*, 58–71.

Aristotle (1999). *Nicomachean ethics* (T. Irwin, Trans.). Hackett. (Original work published 350 BCE.).

Armitage, C. J., & Conner, M. (2000). Attitudinal ambivalence: A test of three key hypotheses. *Personality and Social Psychology Bulletin*, *26*, 1421–1432.

Aronowitz, S., & Lombrozo, T. (2020). Experiential explanation. *Topics in Cognitive Science*, *12*, 1321–1336.

Barber, B. M., & Odean, T. (2000). Trading is hazardous to your wealth: The common stock investment performance of individual investors. *Journal of Finance*, *55*, 773–806.

Barberis, N., Greenwood, R., Jin, L., & Shleifer, A. (2015). X-CAPM: An extrapolative capital asset pricing model. *Journal of Financial Economics*, *115*, 1–24.

Barbey, A. K., & Sloman, S. A. (2007). Base-rate respect: From ecological rationality to dual processes. *Behavioral and Brain Sciences*, *30*, 241–297.

Baron, J. (1992). The effect of normative beliefs on anticipated emotions. *Journal of Personality and Social Psychology*, *63*, 320–330.

Baron, J., & Kemp, S. (2004). Support for trade restrictions, attitudes, and understanding of comparative advantage. *Journal of Economic Psychology*, *25*, 565–580.

Baron, J., & Spranca, M. (1997). Protected values. *Organizational Behavior & Human Decision Processes*, *70*, 1–16.

Barsalou, L. W. (1983). Ad hoc categories. *Memory & Cognition*, *11*, 211–227.

Bartlett, F. C. (1932). *Remembering: An experimental and social study*. Cambridge University Press.

Batteux, E., Bilovich, A., Johnson, S. G. B., & Tuckett, D. (2021b). Negative consequences of failing to communicate uncertainties during a pandemic: an online randomised controlled trial on COVID-19 vaccines. *BMJ Open*, *12*, e051352.

Batteux, E., Khon, Z., Bilovich, A., Johnson, S. G. B., & Tuckett, D. (in press). When do consumers favor overly precise information about investment returns? *Journal of Experimental Psychology: Applied*.

Batteux, E., Bilovich, A., Johnson, S. G. B., & Tuckett, D. (2021a). When certainty backfires: The effects of unwarranted certainty on consumer loyalty. *UCL Working Paper*.

Baumeister, R. F., & Masicampo, E. J. (2010). Conscious thought is for facilitating social and cultural interactions: How mental simulations serve the animal–culture interface. *Psychological Review*, *117*, 945–971.

Beach, L. R. (1998). *Image theory: Theoretical and empirical foundations*. Erlbaum.

Beach, L. R. (2010). *The psychology of narrative thought: How the stories we tell ourselves shape our lives*. Xlibris.

Beach, L. R. (2020). Scenarios as narratives. *Futures & Foresight Science*, *3*, e58.

Beach, L. R., Bissell, B. L., & Wise, J. A. (2016). *A new theory of mind: The theory of narrative thought*. Cambridge Scholars Publishing.

Beckert, J. (2016). *Imagined futures: Fictional expectations and capitalist dynamics*. Harvard University Press.

Bentham, J. (1907/1789). *An introduction to the principles of morals and legislation*. Clarendon Press.

Berger, J. (2013). *Contagious: Why things catch on*. Simon & Schuster.

Berlyne, D. E. (1960). *Conflict, arousal, and curiosity*. McGraw-Hill.

Bhattacharjee, A., Dana, J., & Baron, J. (2017). Anti-profit beliefs: How people neglect the societal benefits of profit. *Journal of Personality and Social Psychology*, *113*, 671–696.

Bilovich, A., Johnson, S. G. B., & Tuckett, D. (2020). Perceived argument strength mediates the influence of emotions on confidence. *UCL Working Paper*.

Bonawitz, E. B., & Lombrozo, T. (2012). Occam's rattle: Children's use of simplicity and probability to constrain inference. *Developmental Psychology*, *48*, 1156–1164.

Bott, L., & Heit, E. (2004). Nonmonotonic extrapolation in function learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 38–50.

Bowers, J. S., & Davis, C. J. (2012). Bayesian just-so stories in psychology and neuroscience. *Psychological Bulletin*, *138*, 389–414.

Boyd, B. (2009). *On the origin of stories: Evolution, cognition, and fiction*. Belknap.

Boyd, B. (2018). The evolution of stories: From mimesis to language, from fact to fiction. *Wiley Interdisciplinary Reviews Cognitive Science*, *9*, e1444.

Boyd, R., & Richerson, P. (1985). *Culture and evolutionary process*. University of Chicago Press.

Boyd, R., Richerson, P. J., & Henrich, J. (2011). The cultural niche: Why social learning is essential for human adaptation. *Proceedings of the National Academy of Sciences*, *108*, 10918–10925.

Boyer, P., & Petersen, M. B. (2018). Folk-economic beliefs: An evolutionary cognitive model. *Behavioral and Brain Sciences*, *41*, e158.

Bramley, N. R., Gerstenberg, T., Mayrhofer, R., & Lagnado, D. A. (2018). Time in causal structure learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*, 1880–1910.

Bruner, J. (1986). *Actual minds, possible worlds*. Harvard University Press.

Buehner, M. J., & May, J. (2002). Knowledge mediates the timeframe of covariation assessment in human causal induction. *Thinking and Reasoning*, *8*, 269–295.

Camerer, C., & Weber, M. (1992). Recent developments in modeling preferences: Uncertainty and ambiguity. *Journal of Risk and Uncertainty*, *5*, 325–370.

Caplan, B. (2007). *The myth of the rational voter: Why democracies choose bad policies*. Princeton University Press.

Carey, S. (2009a). *The origin of concepts*. Oxford University Press.

Carey, S. (2009b). Where our number concepts come from. *Journal of Philosophy*, *106*, 220–254.

Carlin, W., & Soskice, D. (2018). Stagnant productivity and low unemployment: Stuck in a Keynesian equilibrium. *Oxford Review of Economic Policy*, *34*, 169–194.

Carlson, J. A., & Parkin, M. (1975). Inflation expectations. *Economica*, *42*, 123–138.

Chang, C. (2008). Increasing mental health literacy via narrative advertising. *Journal of Health Communication*, *13*, 37–55.

Chater, N., & Loewenstein, G. (2016). The under-appreciated drive for sense-making. *Journal of Economic Behavior & Organization*, *126*, 137–154.

Chong, K., & Tuckett, D. (2015). Constructing conviction through action and narrative: How money managers manage uncertainty and the consequences for financial market functioning. *Socio-Economic Review*, *13*, 1–26.

Chwe, M. S. (2001). *Rational ritual: Culture, coordination, and common knowledge*. Princeton University Press.

Cimpian, A., & Steinberg, O. D. (2014). The inherence heuristic across development: Systematic differences between children's and adults' explanations for everyday facts. *Cognitive Psychology*, *75*, 130–154.

Cushman, F. (2020). Rationalization is rational. *Behavioral and Brain Sciences*, *43*, e28.

Cutler, D. M., Poterba, J. M., & Summers, L. H. (1991). Speculative dynamics. *Review of Economic Studies*, *58*, 529–546.

Damasio, A. (1994). *Descartes' error: Emotion, reason, and the human brain*. Putnam.

Dawkins, R. (1976). *The selfish gene*. Oxford University Press.

Dawson, C., & Johnson, S. G. B. (2021). *Dread aversion and economic preferences*. Available at PsyArXiv.

De Bondt, W. F. M. (1993). Betting on trends: Intuitive forecasts of financial risk and return. *International Journal of Forecasting*, *9*, 355–371.

De Freitas, J., & Johnson, S. G. B. (2018). Optimality bias in moral judgment. *Journal of Experimental Social Psychology*, *79*, 149–163.

DeLosh, E. L., Busemeyer, J. R., & McDaniel, M. A. (1997). Extrapolation: The sine qua non for abstraction in function learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *23*, 968–986.

De Wit, J. B. F., Das, E., & Vet, R. (2008). What works best: Objective statistics or a personal testimonial? An assessment of the persuasive effects of different types of message evidence on risk perception. *Health Psychology*, *27*, 110–115.

Doherty, M. E., Chadwick, R., Caravan, H., Barr, D., & Mynatt, C. R. (1996). On people's understanding of the diagnostic implications of probabilistic data. *Memory & Cognition*, *24*, 644–654.

Donald, M. (1991). *Origins of the modern mind: Three stages in the evolution of culture and cognition*. Harvard University Press.

Du, N., & Budescu, D. V. (2005). The effects of imprecise probabilities and outcomes in evaluating investment options. *Management Science*, *51*, 1791–1803.

Dunbar, R. (1996). *Grooming, gossip, and the evolution of language*. Harvard University Press.

Einhorn, H. J., & Hogarth, R. M. (1986). Judging probable cause. *Psychological Bulletin*, *99*, 3–19.

Eliaz, K., & Spiegler, R. (2018). A model of competing narratives. *Available at arXiv*.

Elliot, A. J. (2006). The hierarchical model of approach-avoidance motivation. *Motivation and Emotion*, *30*, 111–116.

Ellsberg, D. (1961). Risk, ambiguity, and the Savage axioms. *Quarterly Journal of Economics*, *75*, 643–669.

Escalas, J. E. (2004). Imagine yourself in the product: Mental simulation, narrative transportation, and persuasion. *Journal of Advertising*, *33*, 37–48.

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics*, *114*, 817–868.

Fenton O'Creevy, M., & Tuckett, D. (2022). Selecting futures: The role of conviction, narratives, ambivalence, and constructive doubt. *Futures & Foresight Science*, *4*, e111.

Fernbach, P. M., Darlow, A., & Sloman, S. A. (2010). Neglect of alternative causes in predictive but not diagnostic reasoning. *Psychological Science*, *21*, 329–336.

Festinger, L. (1962). Cognitive dissonance. *Scientific American*, *207*, 93–106.

Fishbach, A., & Dhar, R. (2007). Dynamics of goal-based choice. In C. P. Haugtvedt, P. M. Herr, & F. R. Kardes (Eds.), *Handbook of consumer psychology* (pp. 611–637). Psychology Press.

Fisher, M., & Keil, F. C. (2018). The binary bias: A systematic distortion in the integration of information. *Psychological Science*, *29*, 1846–1858.

Fodor, J. A., & Pylyshyn, Z. (1981). How direct is visual perception? Some reflections on Gibson's "ecological approach". *Cognition*, *9*, 139–196.

Forbus, K. D. (1984). Qualitative process theory. *Artificial Intelligence*, *24*, 85–168.

Ford, M. E. (1992). *Motivating humans: Goals, emotions, and personal agency beliefs*. Sage.

Frijda, N. H. (1988). The laws of emotion. *American Psychologist*, *43*, 449–458.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, *7*, 155–170.

Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naïve theory of rational action. *Trends in Cognitive Sciences*, *7*, 287–292.

Gigerenzer, G. (2008). Why heuristics work. *Perspectives on Psychological Science*, *3*, 20–29.

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, *103*, 650–669.

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review*, *102*, 684–704.

Glazer, A., & Konrad, K. A. (1996). A signaling explanation for charity. *American Economic Review*, *86*, 1019–1028.

Goodman, N. (1955). *Fact, fiction, and forecast*. Harvard University Press.

Gopnik, A. (1998). Explanation as orgasm. *Minds and Machines*, *8*, 101–118.

Gopnik, A., Glymour, C., Sobel, D. M., Schulz, L. E., Kushnir, T., & Danks, D. (2004). A theory of causal learning in children: Causal maps and Bayes nets. *Psychological Review*, *111*, 3–32.

Graesser, A. C., Singer, M., & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review*, *101*, 371–395.

Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology*, *79*, 701–721.

Greene, J. D., Sommerville, R. B., Nystrom, L. E., Darley, J. M., & Cohen, J. D. (2001). An fMRI investigation of emotional engagement in moral judgment. *Science*, *293*, 2105–2108.

Greene, M. R., & Fei-Fei, L. (2014). Visual categorization is automatic and obligatory: Evidence from Stroop-like paradigm. *Journal of Vision*, *14*, 1–11.

Gregory, R. (1970). *The intelligent eye*. Weidenfeld and Nicolson.

Griskevicius, V., Tybur, J. M., & van den Bergh, B. (2010). Going green to be seen: Status, reputation, and conspicuous conservation. *Journal of Personality and Social Psychology*, *98*, 392–404.

Gross, J. J. (1998). The emerging field of emotion regulation: An integrative review. *Review of General Psychology*, *2*, 271–299.

Grossmann, I., Eibach, R. P., Koyama, J., & Sahi, Q. B. (2020). Folk standards of sound judgment: Rationality versus reasonableness. *Science Advances*, *6*, eaaz0289.

Grossmann, I., Oakes, H., & Santos, H. C. (2019). Wise reasoning benefits from emodiversity, irrespective of emotional intensity. *Journal of Experimental Psychology: General*, *148*, 805–823.

Hagmayer, Y., & Sloman, S. A. (2009). Decision makers conceive of their choices as interventions. *Journal of Experimental Psychology: General*, *138*, 22–38.

Hahn, K. H. Y., & Bhadun, G. (2021). Mask up: Exploring cross-cultural influences on mask-making behavior during the COVID-19 pandemic. *Clothing and Textiles Research Journal*, *39*, 297–313.

Hamby, A., Brinberg, D., & Daniloski, K. (2017). Reflecting on the journey: Mechanisms in narrative persuasion. *Journal of Consumer Psychology*, *27*, 11–22.

Hastie, R., & Pennington, N. (2000). Explanation-based decision making. In T. Connolly, H. R. Arkes, & K. R. Hammond (Eds.), *Judgment and decision making: An interdisciplinary reader* (pp. 212–228). Cambridge University Press.

Hayek, F. A. (1958). Freedom, reason, and tradition. *Ethics*, *68*, 229–245.

Hayes, B. K., & Chen, T. J. (2008). Clinical expertise and reasoning with uncertain categories. *Psychonomic Bulletin & Review*, *15*, 1002–1007.

Hegarty, M. (2004). Mechanical reasoning by mental simulation. *Trends in Cognitive Sciences*, *8*, 280–285.

Henrich, J. (2018). *The secret of our success: How culture is driving human evolution, domesticating our species, and making us smarter*. Princeton University Press.

Henrich, J. (2020). *The WEIRDest people in the world: How the West became psychologically peculiar and particularly prosperous*. Penguin.

Hirst, W., Yamashiro, J. K., & Coman, A. (2018). Collective memory from a psychological perspective. *Trends in Cognitive Sciences*, *22*, 438–451.

Holyoak, K. J., Lee, H. S., & Lu, H. (2010). Analogical and category-based inference: A theoretical integration with Bayesian causal models. *Journal of Experimental Psychology: General*, *139*, 702–727.

Hommes, C. (2011). The heterogeneous expectations hypothesis: Some evidence from the lab. *Journal of Economic Dynamics and Control*, *35*, 1–24.

Hommes, C., Sonnemans, J., Tuinstra, J., & van de Velden, H. (2005). Coordination of expectations in asset pricing experiments. *Review of Financial Studies*, *18*, 955–980.

Horne, Z., Muradoglu, M., & Cimpian, A. (2019). Explanation as a cognitive process. *Trends in Cognitive Sciences*, *23*, 187–199.

Jara-Ettinger, J., Gweon, H., Schulz, L. E., & Tenenbaum, J. B. (2016). The naïve utility calculus: Computational principles underlying commonsense psychology. *Trends in Cognitive Sciences*, *20*, 589–604.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance*, *48*, 65–91.

Johnson, M. K., & Sherman, S. J. (1990). Constructing and reconstructing the past and future in the present. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (Vol. *2*, pp. 482–526). Guilford Press.

Johnson, S. G. B., & Hill, F. (2017). Belief digitization in economic prediction. In A. Gunzelmann, A. Howes, T. Tenbrink, & E. J. Davelaar (Eds.), *Proceedings of the 39th annual conference of the Cognitive Science Society* (pp. 2313–2319). Cognitive Science Society.

Johnson, S. G. B., Matiashvili, T., & Tuckett, D. (2019a). Explaining the past, predicting the future: How attributions for past price changes affect price expectations. *Working Paper*.

Johnson, S. G. B., & Tuckett, D. (2022). Narrative expectations in financial forecasting. *Journal of Behavioral Decision Making*, *35*, e2245.

Johnson, S. G. B. (2016). Cognition as sense-making. Unpublished doctoral dissertation, Yale University, New Haven, CT.

Johnson, S. G. B. (2019). Toward a cognitive science of markets: Economic agents as sense-makers. *Economics*, *13*, 49.

Johnson, S. G. B. (2020). Dimensions of altruism: Do evaluations of charitable behavior track prosocial benefit or personal sacrifice? *Available at PsyArXiv*.

Johnson, S. G. B., & Ahn, J. (2021). Principles of moral accounting: How our intuitive moral sense balances rights and wrongs. *Cognition*, *206*, 104467.

Johnson, S. G. B., & Ahn, W. (2015). Causal networks or causal islands? The representation of mechanisms and the transitivity of causal judgment. *Cognitive Science*, *39*, 1468–1503.

Johnson, S. G. B., & Ahn, W. (2017). Causal mechanisms. In M. R. Waldmann (Ed.), *Oxford handbook of causal reasoning* (pp. 127–146). Oxford University Press.

Johnson, S. G. B., Jin, A., & Keil, F. C. (2014). Simplicity and goodness-of-fit in explanation: The case of intuitive curve-fitting. In P. Bello, M. Guarini, M. McShane & B. Scassellati (Eds.), *Proceedings of the 36th annual conference of the Cognitive Science Society* (pp. 701–706). Cognitive Science Society.

Johnson, S. G. B., & Keil, F. C. (2014). Causal inference and the hierarchical structure of experience. *Journal of Experimental Psychology: General*, *143*, 2223–2241.

Johnson, S. G. B., Kim, H. S., & Keil, F. C. (2016a). Explanatory biases in social categorization. In A. Papafragou, D. Grodner, D. Mirman, & J. C. Trueswell (Eds.), *Proceedings of the 38th annual conference of the Cognitive Science Society* (pp. 776–781). Cognitive Science Society.

Johnson, S. G. B., Matiashvili, T., & Tuckett, D. (2019b). Expectations based on past price patterns: An experimental study. *UCL Working Paper*.

Johnson, S. G. B., Merchant, T., & Keil, F. C. (2015a). Argument scope in inductive reasoning: Evidence for an abductive account of induction. In D. C. Noelle, R. Dale, A. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings, & P. P. Maglio (Eds.), *Proceedings of the 37th annual conference of the Cognitive Science Society* (pp. 1015–1020). Cognitive Science Society.

Johnson, S. G. B., Merchant, T., & Keil, F. C. (2020). Belief digitization: Do we treat uncertainty as probabilities or as bits? *Journal of Experimental Psychology: General*, *149*, 1417–1434.

Johnson, S. G. B., Murphy, G. L., Rodrigues, M., & Keil, F. C. (2019). Predictions from uncertain moral character. In A. K. Goel, C. M. Seifert, & C. Freksa (Eds.), *Proceedings of the 41st annual conference of the Cognitive Science Society* (pp. 506–512). Cognitive Science Society.

Johnson, S. G. B., & Park, S. Y. (2021). Moral signaling through donations of money and time. *Organizational Behavior & Human Decision Processes*, *165*, 183–196.

Johnson, S. G. B., Rajeev-Kumar, G., & Keil, F. C. (2015b). Belief utility as an explanatory virtue. In D. C. Noelle, R. Dale, A. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings, & P. P. Maglio (Eds.), *Proceedings of the 37th annual conference of the Cognitive Science Society* (pp. 1009–1014). Cognitive Science Society.

Johnson, S. G. B., Rajeev-Kumar, G., & Keil, F. C. (2016b). Sense-making under ignorance. *Cognitive Psychology*, *89*, 39–70.

Johnson, S. G. B., & Rips, L. J. (2015). Do the right thing: The assumption of optimality in lay decision theory and causal judgment. *Cognitive Psychology*, *77*, 42–76.

Johnson, S. G. B., Rodrigues, M., & Tuckett, D. (2021). Moral tribalism and its discontents: How intuitive theories of ethics shape consumers' deference to experts. *Journal of Behavioral Decision Making*, *34*, 47–65.

Johnson, S. G. B., & Steinerberger, S. (2019). Intuitions about mathematical beauty: A case study in the aesthetic experience of ideas. *Cognition*, *189*, 242–259.

Johnson, S. G. B., Valenti, J. J., & Keil, F. C. (2019). Simplicity and complexity preferences in causal explanation: An opponent heuristic account. *Cognitive Psychology*, *113*, 101222.

Johnson, S. G. B., Zhang, J., & Keil, F. C. (2018). Psychological underpinnings of zero-sum thinking. In T. T. Rogers, M. Rau, X. Zhu, & C. W. Kalish (Eds.), *Proceedings* *of the 40th annual conference of the Cognitive Science Society* (pp. 566–571). Cognitive Science Society.

Johnson, S. G. B., Zhang, J., & Keil, F. C. (2019). Consumers' beliefs about the effects of trade. *Available at SSRN*.

Johnston, A. M., Johnson, S. G. B., Koven, M. L., & Keil, F. C. (2017). Little Bayesians or little Einsteins? Probability and explanatory virtue in children's inferences. *Developmental Science*, *20*, e12483.

Johnston, A. M., Sheskin, M., Johnson, S. G. B., & Keil, F. C. (2018). Preferences for explanation generality develop early in biology, but not physics. *Child Development*, *89*, 1110–1119.

Jones, M., & Love, B. C. (2011). Bayesian fundamentalism or enlightenment? On the explanatory status and theoretical contributions of Bayesian models of cognition. *Behavioral and Brain Sciences*, *34*, 169–188.

Jordan, J. J., Hoffman, M., Bloom, P., & Rand, D. G. (2016). Third-party punishment as a costly signal of trustworthiness. *Nature*, *530*, 473–476.

Kabiri, A., James, H., Landon-Lane, J., Tuckett, D., & Nyman, R. (2023). The role of sentiment in the US economy: 1920 to 1934. *Economic History Review*, *76*, 3–30.

Kahneman, D. (2002). Maps of bounded rationality: Psychology for behavioral economics. *American Economic Review*, *93*, 1449–1475.

Kahneman, D., & Tversky, A. (1982). The simulation heuristic. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 201–208). Cambridge University Press.

Kant, I. (2002). *Groundwork for the metaphysics of morals* (A. Zweig, Trans.). Oxford University Press. (Original work published 1796).

Kay, J., & King, M. (2020). *Radical uncertainty: Decision-making for an unknowable future*. Bridge Street Press.

Keynes, J. M. (1936). *The general theory of employment, interest, and money*. Palgrave Macmillan.

Khemlani, S. S., Sussman, A. B., & Oppenheimer, D. M. (2011). *Harry Potter* and the sorcerer's scope: Latent scope biases in explanatory reasoning. *Memory & Cognition*, *39*, 527–535.

Kintsch, W., Mandel, T. S., & Kozminsky, E. (1977). Summarizing scrambled stories. *Memory and Cognition*, *5*, 547–552.

Klein, G. (1998). *Sources of power: How people make decisions*. MIT Press.

Knight, F. (1921). *Risk, uncertainty, and profit*. Houghton-Mifflin.

Knobe, J. (2010). Person as scientist, person as moralist. *Behavioral and Brain Sciences*, *33*, 315–329.

Koehler, J. J. (1996). The base rate fallacy reconsidered: Descriptive, normative, and methodological challenges. *Behavioral and Brain Sciences*, *19*, 1–53.

Kominsky, J. F., Phillips, J., Gerstenberg, T., Lagnado, D., & Knobe, J. (2015). Causal superseding. *Cognition*, *137*, 196–209.

Kosslyn, S. M. (1975). Information representation in visual images. *Cognitive Psychology*, *7*, 341–370.

Krause, R. J., & Rucker, D. D. (2020). Strategic storytelling: When narratives help versus hurt the persuasive power of facts. *Personality and Social Psychology Bulletin*, *46*, 216–227.

Krynski, T. R., & Tenenbaum, J. B. (2007). The role of causality in judgment under uncertainty. *Journal of Experimental Psychology: General*, *136*, 430–450.

Lagnado, D. A., & Shanks, D. R. (2003). The influence of hierarchy on probability judgment. *Cognition*, *89*, 157–178.

Lagnado, D. A., & Sloman, S. A. (2006). Time as a guide to cause. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 451–460.

Lake, B. M., Ullman, T. D., Tenenbaum, J. B., & Gershman, S. J. (2017). Building machines that learn and think like people. *Behavioral and Brain Sciences*, *40*, e253.

Lavoie, M. (1994). A post Keynesian approach to consumer choice. *Journal of Post Keynesian Economics*, *16*, 539–562.

Lazarus, R. S. (1991). *Emotion and adaptation*. Oxford University Press.

LeBoeuf, R. A., & Norton, M. I. (2012). Consequence–cause matching: Looking to the consequences of events to infer their causes. *Journal of Consumer Research*, *39*, 128–141.

Leiser, D., & Aroch, R. (2009). Lay understanding of macroeconomic causation: The good-begets-good heuristic. *Applied Psychology*, *58*, 370–384.

Leiser, D., & Shemesh, Y. (2018). *How we misunderstand economics and why it matters: The psychology of bias, distortion, and conspiracy*. Routledge.

Lerner, J. S., & Keltner, D. (2000). Beyond valence: Toward a model of emotion-specific influences on judgement and choice. *Cognition and Emotion*, *14*, 473–493.

Lerner, J. S., & Keltner, D. (2001). Fear, anger, and risk. *Journal of Personality and Social Psychology*, *81*, 146–159.

Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the effects of accountability. *Psychological Bulletin*, *125*, 255–275.

Lewin, K. (1935). *A dynamic theory of personality*. McGraw-Hill.

Lieder, F., & Griffiths, T. L. (2020). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences*, *43*, e1.

Lipton, P. (2004). *Inference to the best explanation* (2nd ed.). Routledge.

Loewenstein, G. (1987). Anticipation and the valuation of delayed consumption. *The Economic Journal*, *97*, 666–684.

Loewenstein, G. F., Thompson, L., & Bazerman, M. H. (1989). Social utility and decision making in interpersonal contexts. *Journal of Personality and Social Psychology*, *57*, 426–441.

Loewenstein, G. F., Weber, E. U., Hsee, C. K., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin*, *127*, 267–286.

Lombrozo, T. (2007). Simplicity and probability in causal explanation. *Cognitive Psychology*, *55*, 232–257.

Lombrozo, T. (2016). Explanatory preferences shape learning and inference. *Trends in Cognitive Sciences*, *20*, 748–759.

Loomes, G., & Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *The Economic Journal*, *92*, 805–824.

Malle, B. F. (1999). How people explain behavior: A new theoretical framework. *Personality and Social Psychology Review*, *3*, 23–48.

Mandler, J. M., & Johnson, N. S. (1977). Remembrance of things parsed: Story structure and recall. *Cognitive Psychology*, *9*, 111–151.

Mar, R. A., Oatley, K., Dikic, M., & Mullin, J. (2010). Emotion and narrative fiction: Interactive influences before, during, and after reading. *Cognition and Emotion*, *25*, 818–833.

Marcus, G. F., & Davis, E. (2013). How robust are probabilistic models of higher-level cognition? *Psychological Science*, *24*, 2351–2360.

Markman, K. D., Gavanski, I., Sherman, S. J., & McMullen, M. N. (1993). The mental simulation of better and worse possible worlds. *Journal of Experimental Social Psychology*, *29*, 87–109.

Marks, J., Copland, E., Loh, E., Sunstein, C. R., & Sharto, T. (2019). Epistemic spillovers: Learning others' political views reduces the ability to assess and use their expertise in nonpolitical domains. *Cognition*, *188*, 74–84.

Marr, D. (1982). *Vision*. Freeman.

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, *50*, 370–396.

Mercier, H., & Sperber, D. (2017). *The enigma of reason*. Harvard University Press.

Mikhail, J. (2007). Universal moral grammar: Theory, evidence, and the future. *Trends in Cognitive Sciences*, *11*, 143–152.

Miller, G. A. (1956) The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*, 81–97.

Mitchell, J. P., Banaji, M. R., & Macrae, C. N. (2005). The link between social cognition and self-referential thought in the medial prefrontal cortex. *Journal of Cognitive Neuroscience*, *17*, 1306–1315.

Moors, A., & De Houwer, J. (2010). Automatic appraisal of motivational valence: Motivational affective priming and Simon effects. *Cognition and Emotion*, *15*, 749–766.

Mulligan, E. J., & Hastie, R. (2005). Explanations determine the impact of information on financial investment judgments. *Journal of Behavioral Decision Making*, *18*, 145–156.

Murphy, G. L., & Medin, D. L. (1985). The role of theories in conceptual coherence. *Psychological Review*, *92*, 289–316.

Murphy, G. L., & Ross, B. H. (1994). Predictions from uncertain categorizations. *Cognitive Psychology*, *27*, 148–193.

Nichols, S. (2002). Norms with feeling: Towards a psychological account of moral judgment. *Cognition*, *84*, 221–236.

Nichols, S., & Knobe, J. (2007). Moral responsibility and determinism: The cognitive science of folk intuitions. *Noûs*, *41*, 663–685.

Norenzayan, A., Atran, S., Faulkner, J., & Schaller, M. (2006). Memory and mystery: The cultural selection of minimally counterintuitive narratives. *Cognitive Science*, *30*, 531–553.

Nussbaum, M. C. (2001). *Upheavals of thought: The intelligence of emotions*. Cambridge University Press.

Nyman, R., Kapadia, S., & Tuckett, D. (2021). News and narratives in financial systems: Exploiting big data for systemic risk assessment. *Journal of Economic Dynamics & Control*, *127*, 104119.

Oatley, K., & Johnson-Laird, P. N. (1987). Towards a cognitive theory of emotions. *Cognition and Emotion*, *1*, 29–50.

Panksepp, J. (1998). The periconscious substrates of consciousness: Affective states and the evolutionary origins of the self. *Journal of Consciousness Studies*, *5*, 566–582.

Paul, L. A. (2014). *Transformative experience*. Oxford University Press.

Payne, J. W., Bettman, J. R., & Johnson, E. J. (1988). Adaptive strategy selection in decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 534–552.

Pearl, J. (2000). *Causality: Models, reasoning, and inference*. Cambridge University Press.

Pennington, N., & Hastie, R. (1986). Evidence evaluation in complex decision making. *Journal of Personality and Social Psychology*, *51*, 242–258.

Pennington, N., & Hastie, R. (1988). Explanation-based decision making: Effects of memory structure on judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 521–533.

Pennington, N., & Hastie, R. (1992). Explaining the evidence: Tests of the Story Model for juror decision making. *Journal of Personality and Social Psychology*, *62*, 189–206.

Pennington, N., & Hastie, R. (1993). Reasoning in explanation-based decision making. *Cognition*, *49*, 123–163.

Rand, D. G., & Nowak, M. A. (2013). Human cooperation. *Trends in Cognitive Sciences*, *17*, 413–425.

https://doi.org/10.1017/S0140525X22001157 Published online by Cambridge University Press

Rawls, J. (2001). *Justice as fairness: A restatement* (E. Kelly, Ed.). Harvard University Press.

Raz, J. (1986). *The morality of freedom*. Clarendon Press.

Read, S. J., & Marcus-Newhall, A. (1993). Explanatory coherence in social explanations: A parallel distributed processing account. *Journal of Personality and Social Psychology*, *65*, 429–447.

Richard, R., van der Pligt, J., & de Vries, N. (1996). Anticipated affect and behavioral choice. *Basic and Applied Social Psychology*, *18*, 111–129.

Richardson, S., Dohrenwend, B. S., & Klein, D. (1965). *Interviewing: Its forms and functions*. Basic Books.

Ridley, M. (2020). *How innovation works*. Fourth Estate.

Rips, L. J. (2010). Two causal theories of counterfactual conditionals. *Cognitive Science*, *34*, 175–221.

Rolls E. T. (2014). *Emotion and decision-making explained*. Oxford University Press.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, *8*, 382–439.

Rosner, A., Basieva, I., Barque-Duran, A., Glöckner, A., von Helversen, B., Khrennikov, A., & Pothos, E. M. (2022). Ambivalence in decision making: An eye tracking study. *Cognitive Psychology*, *134*, 101464.

Rothman, N. B., Pratt, M. G., Rees, L., & Vogus, T. J. (2017). Understanding the dual nature of ambivalence: Why and when ambivalence leads to good and bad outcomes. *Academy of Management Annals*, *11*, 33–72.

Rottman, B. M., & Hastie, R. (2014). Reasoning about causal relationships: Inferences on causal networks. *Psychological Bulletin*, *140*, 109–139.

Rottman, B. M., & Keil, F. C. (2012). Causal structure learning over time: Observations and interventions. *Cognitive Psychology*, *64*, 93–125.

Rozenblit, L., & Keil, F. C. (2002). The misunderstood limits of folk science: An illusion of explanatory depth. *Cognitive Science*, *26*, 521–562.

Rozin P., & Fallon, A. E. (1987). A perspective on disgust. *Psychological Review*, *94*, 23–41.

Rubin, D. C. (1995). *Memory in oral traditions: The cognitive psychology of epic, ballads, and counting-out rhymes*. Oxford University Press.

Rucker, D. D., Tormala, Z. L., Pety, R. E., & Briñol, P. (2014). Consumer conviction and commitment: An appraisal-based framework for attitude certainty. *Journal of Consumer Psychology*, *24*, 119–136.

Rumelhart, D. E. (1975). Notes on a schema for stories. In D. G. Bobrow & A. Collins (Eds.), *Representation and understanding: Studies in cognitive science* (pp. 211–236). Academic Press.

Sanborn, A. N., & Chater, N. (2016). Bayesian brains without probabilities. *Trends in Cognitive Sciences*, *20*, 883–893.

Sanborn, A. N., Griffiths, T. L., & Navarro, D. J. (2010). Rational approximations to rational models: Alternative algorithms for category learning. *Psychological Review*, *117*, 1144–1167.

Savage, L. J. (1954). *The foundations of statistics*. Wiley.

Schacter, D. L., & Addis, D. R. (2007). The ghosts of past and future: A memory that works by piecing together bits of the past may be better suited to simulating future events than one that is a store of perfect records. *Nature*, *445*, 27.

Schacter, D. L., Addis, D. R., & Buckner, R. L. (2008). Episodic simulation of future events: Concepts, data, and applications. *Annals of the New York Academy of Sciences*, *1124*, 39–60.

Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals, and understanding: An inquiry into human knowledge structures*. Erlbaum.

Schwarz, N. (1990). Feelings as information: Informational and motivational functions of affective states. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition: Cognitive foundations of social psychology* (Vol. *2*, pp. 527–561). Guilford Press.

Shafir, E., Simonson, I., & Tversky, A. (1993). Reason-based choice. *Cognition*, *49*, 11–36.

Shefrin, H. (2002). *Beyond greed and fear: Understanding behavioral finance and the psychology of investing* (2nd ed.). Oxford University Press.

Shen, F., Ahern, L., & Baker, M. (2014). Stories that count: Influence of news narratives on issue attitudes. *Journalism and Mass Communication Quarterly*, *91*, 98–117.

Shiller, R. J. (1981). Do stock prices move too much to be justified by subsequent changes in dividends? *American Economic Review*, *71*, 421–436.

Shiller, R. J. (2019). *Narrative economics: How stories go viral and drive major economic events*. Princeton University Press.

Shtulman, A. (2017). *Scienceblind: Why our intuitive theories about the world are so often wrong*. Basic Books.

Sibley, W. M. (1953). The rational and the reasonable. *Philosophical Review*, *62*, 554–560.

Silvia, P. J. (2008). Interest – the curious emotion. *Current Directions in Psychological Science*, *17*, 57–60.

Simon, H. A. (1955). A behavioral model of rational choice. *Quarterly Journal of Economics*, *69*, 99–118.

Simon, H. A. (1957). *Models of man*. Wiley.

Sinnott-Armstrong, W. (1985). Moral dilemmas and incomparability. *American Philosophical Quarterly*, *22*, 321–329.

Sloman, S. (2005). *Causal models*. Oxford University Press.

Sloman, S., & Fernbach, P. (2017). *The knowledge illusion: Why we never think alone*. Penguin Random House.

Sloman, S. A., & Hagmayer, Y. (2006). The causal psycho-logic of choice. *Trends in Cognitive Sciences*, *10*, 407–412.

Smelser, N. J. (1998). The rational and the ambivalent in the social sciences. *American Sociological Review*, *63*, 1–16.

Smith, C. A., & Ellsworth, P. C. (1985). Patterns of cognitive appraisal in emotion. *Journal of Personality and Social Psychology*, *48*, 813–838.

Smith, D., Schlaepfer, P., Major, K., Dyble, M., Page, A. E., Thompson, J., … Migliano, B. (2017). Cooperation and the evolution of hunter–gatherer storytelling. *Nature Communications*, *8*, 1853.

Sparrow, B., Liu, J., & Wegner, D. M. (2011). Google effects on memory: Cognitive consequences of having information at our fingertips. *Science*, *333*, 776–778.

Sperber, D. (1996). *Explaining culture: A naturalistic approach*. Blackwell.

Spirtes, P., Glymour, C., & Scheines, R. (1993). *Causation, prediction, and search*. Springer.

Steiger, J. H., & Gettys, C. F. (1972). Best-guess errors in multistage inference. *Journal of Experimental Psychology*, *92*, 1–7.

Steyvers, M., Tenenbaum, J. B., Wagenmakers, E., & Blum, B. (2003). Inferring causal networks from observations and interventions. *Cognitive Science*, *27*, 453–489.

Strickland, B., Silver, I., & Keil, F. C. (2016). The texture of causal construals: Domain-specific biases shape causal inferences from discourse. *Memory & Cognition*, *45*, 442–455.

Suddendorf, T., & Corballis, M. C. (1997). Mental time travel and the evolution of the human mind. *Genetic, Social, and General Psychology Monographs*, *123*, 133–167.

Suddendorf, T., & Corballis, M. C. (2007). The evolution of foresight: What is mental time travel, and is it unique to humans? *Behavioral and Brain Sciences*, *30*, 299–313.

Sunstein, C. R. (2005). Moral heuristics. *Behavioral and Brain Sciences*, *28*, 531–573.

Sussman, A. B., Khemlani, S. S., & Oppenheimer, D. M. (2014). Latent scope bias in categorization. *Journal of Experimental Social Psychology*, *52*, 1–8.

Sussman, A. B., & Oppenheimer, D. M. (2020). The effect of effects on effectiveness: A boon-bane asymmetry. *Cognition*, *199*, 104240.

Taleb, N. N. (2001). *Fooled by randomness: The hidden role of chance in life and in the markets*. Random House.

Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to grow a mind: Statistics, structure, and abstraction. *Science*, *331*, 1379–1385.

Tennie, C., Frith, U., & Frith, C. D. (2010). Reputation management in the age of the world-wide web. *Trends in Cognitive Sciences*, *14*, 482–488.

Tetlock, P. E. (2003). Thinking the unthinkable: Sacred values and taboo cognitions. *Trends in Cognitive Sciences*, *7*, 320–324.

Thagard, P. (1989). Explanatory coherence. *Behavioral and Brain Sciences*, *12*, 435–502.

Thorndyke, P. W. (1977). Cognitive structures in comprehension and memory of narrative discourse. *Cognitive Psychology*, *9*, 77–110.

Todd, P. M., & Gigerenzer, G. (2007). Environments that make us smart: Ecological rationality. *Current Directions in Psychological Science*, *16*, 167–171.

Todorov, A. (2008). Evaluating faces on trustworthiness: An extension of systems for recognition of emotions signaling approach/avoidance behaviors. *Annals of the New York Academy of Sciences*, *1124*, 208–224.

Trope, Y., & Liberman, N. (2003). Temporal construal. *Psychological Review*, *110*, 403–421.

Tuckett, D. (2011). *Minding the markets: An emotional finance view of financial instability*. Palgrave Macmillan.

Tuckett, D. (2012). Financial markets are markets in stories: Some possible advantages of using interviews to supplement existing economic data sources. *Journal of Economic Dynamics and Control*, *36*, 1077–1087.

Tuckett, D. (2017). Why observation of the behaviour of human actors and how they combine within the economy is an important next step. Paper presented at the *Institute of New Economic Thinking: Future of Macroeconomics Conference*, Edinburgh, UK.

Tuckett, D., Boulton, M., & Olson, C. (1985). A new approach to the measurement of patients' understanding of what they are told in medical consultations. *Journal of Health and Social Behavior*, *26*, 27–38.

Tuckett, D., Holmes, D., Pearson, A., & Chaplin, G. (2020). Monetary policy and the management of uncertainty: A narrative approach. *Bank of England Working Paper*.

Tuckett, D. (in press). Feelings, narratives, and mental states: How neuroeconomics can shift the paradigm. In A. Kirman & M. Teschi (Eds.), *The state of mind in economics*. Cambridge University Press.

Tuckett, D., & Nikolic, M. (2017). The role of conviction and narrative in decision-making under radical uncertainty. *Theory & Psychology*, *27*, 501–523.

Tuckett, D., & Nyman, R. (2018). The relative sentiment shift series for tracking the economy. *UCL Working Paper*.

Tuckett, D., Smith, R. E., & Nyman, R. (2014). Tracking phantastic objects: A computer algorithmic investigation of narrative evolution in unstructured data sources. *Social Networks*, *38*, 121–133.

Tuckett, D., & Taffler, R. (2008). Phantastic objects and the financial market's sense of reality: A psychoanalytic contribution to the understanding of stock market instability. *International Journal of Psychoanalysis*, *89*, 389–412.

Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review*, *79*, 281–299.
Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin*, *76*, 105–110.
Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124–1131.
Tversky, A., & Kahneman, D. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*, 263–291.
Tversky, A., & Kahneman, D. (1980). Causal schemas in judgments under uncertainty. In M. Fishbein (Ed.), *Progress in social psychology* (pp. 49–72). Erlbaum.
Uhlmann, E. L., Pizarro, D. A., & Diermeier, D. (2015). A person-centered approach to moral judgment. *Perspectives on Psychological Science*, *10*, 72–81.
Van Laer, T., de Ruyter, K., Visconti, L. M., & Wetzels, M. (2014). The extended transportation-imagery model: A meta-analysis of the antecedents and consequences of consumers' narrative transportation. *Journal of Consumer Research*, *40*, 797–817.
Van Osselaer, S. M. J., & Janiszewski, C. (2012). A goal-based model of product evaluation and choice. *Journal of Consumer Research*, *39*, 260–292.
Vasilyeva, N., Blanchard, T., & Lombrozo, T. (2018). Stable causal relationships are better causal relationships. *Cognitive Science*, *42*, 1265–1296.
Vasilyeva, N., Wilkenfeld, D., & Lombrozo, T. (2017). Contextual utility affects the perceived quality of explanations. *Psychonomic Bulletin and Review*, *24*, 1436–1450.
Volz, K. G., & Gigerenzer, G. (2012). Cognitive processes in decisions under risk are not the same as in decisions under uncertainty. *Frontiers in Neuroscience*, *6*, 105.
Von Helmholtz, H. (2005). *Treatise on physiological optics* (Vol. *III*). Dover. (Original work published 1867).
Von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton University Press.
Walasek, L., & Brown, G. D. A. (2021). Incomparability and incommensurability in choice: No common currency of value? *Warwick Working Paper*.
Wegner, D. M. (1987). Transactive memory: A contemporary analysis of the group mind. In B. Mullen & G. R. Goethals (Eds.), *Theories of group behavior* (pp. 185–208). Springer–Verlag.
Wilson, T. D., & Gilbert, D. T. (2005). Affective forecasting: Knowing what to want. *Current Directions in Psychological Science*, *14*, 131–134.
Woodward, J. (2003). *Making things happen: A theory of causal explanation*. Oxford University Press.
Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin*, *127*, 3–21.
Zeelenberg M., & Pieters, R. (2004). Consequences of regret aversion in real life: The case of the Dutch postcode lottery. *Organizational Behavior and Human Decision Processes*, *93*, 155–168.
Zemla, J. C., Sloman, S., Bechlivanidis, C., & Lagnado, D. A. (2017). Evaluating everyday explanations. *Psychonomic Bulletin & Review*, *24*, 1488–1500.
Zhu, J., & Murphy, G. L. (2013). Influence of emotionally charged information on category-based induction. *PLoS ONE*, *8*, e54286.

# Open Peer Commentary

## Narratives need not end well; nor say it all

Sara Andreetta*, Davide Spalla† and Alessandro Treves

Cognitive Neuroscience, SISSA, 34136 Trieste, Italy.
sara.andreetta@ung.si
davide.spalla@donders.ru.nl
ale@sissa.it
https://people.sissa.it/~ale/limbo.html

doi:10.1017/S0140525X22002655, e83

**Abstract**

To fully embrace situations of radical uncertainty, we argue that the theory should abandon the requirements that narratives, in general, must lead to *affective evaluation*, and that they have to *explain* (and potentially *simulate*) all or even the bulk of the current decisional context. Evidence from studies of incidental learning show that narrative schemata can bias decisions while remaining fragmentary, insufficient for prediction, and devoid of utility values.

Conviction Narrative Theory (CNT) is a welcome broadening of traditional perspectives on decision-making. It can impact a neural computation-based understanding of cognitively intense situations such as economic evaluation, cultural behavior, and social learning, among others. For the latter in particular, the authors highlight the role of causal schemata, memorable analogies, and emotions, as well as of certain constraints on encoding, communication, and action, in order for these narrative fragments to socially propagate.

A core strength of the theory is in its free articulation in several component features. We argue that the theory is weakened, however, by imposing them as *necessary* components. CNT promises a liberation from the artificial intelligence quagmire of decision trees and reward optimization, but the actual liberation hinges on recognizing the fragmentary character of many narratives, and their *a priori* independence from considerations of affective value.

Consider, for example, the features of narratives which make them easier to remember: Their abstracted scripts, or what has been specifically defined as "story grammars" (Rumelhart, 1975). These patterns have facilitated the transmission of knowledge in oral form for centuries (Rubin, 1995), to the point that some scholars have placed the contribution of narrative schemata among the most salient enabling characteristics of human cognition (Ferretti, 2022; Gottschiel, 2012). To serve as a memory facilitator, a narrative schema obviously does not have to be exact; nor does it have to be complete, or tinged with a reward value.

An example is provided by metric structures in poetry that can be regarded as non-verbal narrative schemata. While fragmentary, their constrained and repetitive nature makes them a suitable model for a quantitative assessment of the contribution of an underlying "narrative" to memory-related decision-making. A recent study of ours (Andreetta, Soldatkina, Boboeva, & Treves, 2021) focuses on the role of metric schemata in remembering poetry: Activating such schemata was shown to help, incrementally, in retrieving nonsense words from a previously heard meaningless "poem," in the absence of any affective value or conventional narration. Therefore, a simple form of decision-making (choosing the previously heard non-word from a choice of three) can be facilitated by the "narrative schema" even if the latter is unrelated to the former (all three options would fit the metric pattern). An interpretation is that the metric narration pushes the flow of neural activation forward, enhancing associative retrieval dynamics.

A mechanistic network model of such neural dynamics, streamlined to mathematically tractable form, has been analyzed by Spalla et al. (2021).

This study shows how narrative fragments can be acquired incidentally through Hebbian self-organization, without any notion of utility, and stored as dynamical attractors in simple recurrent

*Current address: School of Humanities, Univerza v Novi Gorici, Vipavska 13, SI-5000 Nova Gorica, Slovenia
†Current address: Postbus 9010 6500 GL Nijmegen, the Netherlands

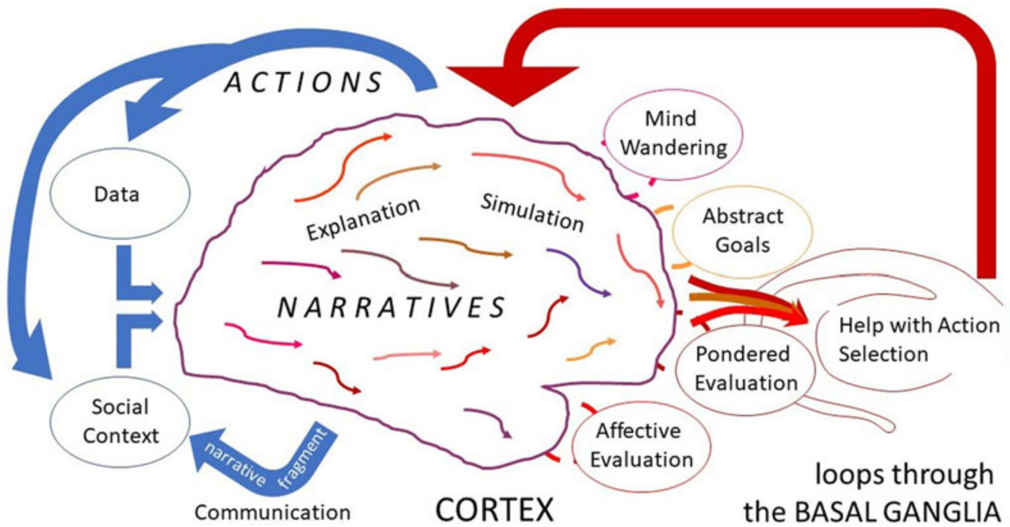


**Figure 1.** (Andreetta et al.) A proposed modification of Figure 2 in the target article, separating internal (warm colors) and external (blue) processes.

networks, ubiquitously in the cortex (as indicated schematically in Fig. 1). In addition, their unfolding in time (the feature that makes such attractors dynamical, and thus suitable to represent fragments of narratives in the brain) does not consume extra storage resources, as shown by a mathematical analysis of the network model. It even pays off for the stored attractors to represent dynamical narratives rather than static scenes.

Finally, the relevance of narrative fragments in shaping brain activity finds experimental support in a recent study by Zheng et al. (2022), indicating that neurons in the human medial temporal lobe "detect cognitive boundaries" in episodic memories.

Therefore, we believe that CNT can be extended from a purely cognitive domain to that of cortical operations, to inform neuroscience research and bridge the socially relevant gap between semi-rational decision-making and the computational constraints that (loosely) bound our thinking.

**Acknowledgement.** None.

**Financial support.** Research funded by Human Frontier Science Program grant RGP0057/2016.

**Competing interest.** None.

## References

Andreetta, S., Soldatkina, O., Boboeva, V., & Treves, A. (2021). In poetry, if meter has to help memory, it takes its time. *Open Research Europe*, *1*, 59. https://doi.org/10.12688/openreseurope.13663.1
Ferretti, F. (2022). *L'istinto persuasivo*. Carocci Editore.
Gottschiel, J. (2012). *The storytelling animal: How stories make us human*. Houghton Mifflin Harcourt.
Rubin, D. C. (1995). *Memory in oral traditions: The cognitive psychology of epic, ballads, and counting-out rhymes*. Oxford University Press on Demand.
Rumelhart, D. E. (1975). Notes on a schema for stories. In D. G. Bobrow & A. Collins (Eds.), *Representation and understanding: Studies in Cognitive Science* (pp. 211–236). Academic Press.
Spalla, D., Cornacchia, I. M., & Treves, A. (2021). Continuous attractors for dynamic memories. *Elife*, *10*(2021), e69499.
Zheng, J., Schjetnan, A. G., Yebra, M., Gomes, B. A., Mosher, C. P., Kalia, S. K., … Rutishauser, U. (2022). Neurons detect cognitive boundaries to structure episodic memories in humans. *Nature Neuroscience*, *25*(3), 358–368.

# Conviction Narrative Theory and the Theory of Narrative Thought

Lee Roy Beach[a] and James A. Wise[b]

[a]Business and Psychology, University of Arizona, Tucson, AZ 85715, USA and and [b]Richland, WA 99352, USA.
leerbeach@aol.com
jamesawise@me.com

doi:10.1017/S0140525X22002795, e84

**Abstract**

Conviction Narrative Theory bears a close resemblance to the Theory of Narrative Thought, although the two were designed to address different questions. In this commentary, we detail some of the more pronounced similarities and differences and suggest that resolving the latter could produce a third theory of narrative cognition that is superior to either of these two.

It is encouraging when the conclusions reached by others on a topic of mutual interest are similar to our own. This is the case with Conviction Narrative Theory (CNT) and our own theory of narrative cognition, the Theory of Narrative Thought, or TNT (e.g., Beach, 2009, 2010, 2018, 2019; Beach, Bissell, & Wise, 2016; Beach & Wise, 2022).

Although CNT is couched as solving a particular problem in decision-making, it can be interpreted as a general theory of narrative cognition, which is what TNT is. This results in two theories, CNT and TNT, with many similarities and some important differences. It seems to us that reconciliation of those differences could produce a single theory of narrative cognition that improves on both of them.

The reader is familiar with CNT, but perhaps not TNT. Briefly, TNT proposes that cognition is ultimately about what will happen in the future and how to avoid or take advantage of it. It posits that the brain structures experience such that events in the past caused present events, which will cause events in the *expected future*. This (implied) expected future is evaluated for potential *threats* – dangers as well as potential loss of good things or loss of opportunities to attain them. Well-being, sometimes even survival, requires identification and mitigation of expected threats before the future arrives and the threats become a reality. Mitigation requires *action* and the question is which action will produce a less threatening future than that which is expected? TNT offers a simple, emotion-based, decision model, called the *discrepancy test*, for both threat appraisal and action selection.

Temporal/causal structuring of experience results in a *narrative* (e.g., Atkinson, 1978; Carroll, 2001). Your own structured experience is called your *prime narrative* (PN), and it constitutes your reality. The PN's content is too great for practical use, but context guides abridgment of pertinent parts of it, called *derived narratives*, for communicating with oneself (thinking) and with other people. What is thought and/or communicated is part of your experience and is therefore incorporated back into your PN. So too, when the future turns out to be different from what was expected, the discrepancies are incorporated back into your PN. Both of these update your PN and increase its *consistency*. The more consistent it or any derived narratives are, the more certain you feel they and what they imply about the future is true and a solid basis for action.

When a derived narrative is shared with other people, it becomes public. Then, others may make it more complex or abstract (beyond one person's private experience) adding to, honing it, and applying it more broadly. Part of the genius of humanity is the collective, cultural elaboration of derived narratives into science, government, religion, etc., all of which exist, in some form, to mitigate threat.

Turning to CNT, its elements are described in the target article's Table 1, which we will use to compare CNT and TNT. The first section, *Context*, describes the decision problem CNT is intended to address; a problem, incidentally, that not everyone regards as a problem (e.g., Phillips, 1970). Be that as it may, the theory's potential goes beyond this limited problem because it can be read as a more general theory of cognition.

Moving to the second section, *Representation*: *Narrative* is defined as incorporating causal, temporal, analogical, and valence information that serves to explain data, imagine and evaluate possible futures, and motivate action. With the exception of analogy, this definition is compatible with TNT's. Analogy bears more attention in TNT. There is nothing like the PN in CNT. *Imagined Futures* is compatible with TNT up to the proviso that they are a response to a contemplated choice because neither CNT nor TNT need be restricted to decision-making. *Narrative Fragments* and *Shared Narratives* correspond to TNT's derived narratives as they stand.

The third section of the table, *Processes*, is more difficult to evaluate because of CNT's focus upon that narrow decision problem. *Explanation*, in the active sense intended by CNT, is not part of TNT and certainly its dependence on heuristics is not either. In TNT, explanation, understanding, comprehension, etc., are all manifestations of the internal consistency of the PN, or a narrative derived from it. *Simulation* generates an imagined future. In TNT, the future simply is the causal implications of what led up to now. Alternative futures can be imagined, but *simulation* may not be the right term for how this happens (see Beach & Wise, 2022). *Affective Evaluation* is more clearly detailed in TNT's Discrepancy Test (ibid.), but the ideas are essentially the same. *Communication* is the same in both theories. There is (at least) one place where CNT far outpaces TNT. CNT's description of causal rules that, together with time, define narratives is more thorough and solid that what we have advanced in TNT.

There are other similarities and differences, some of which simply result from the two theories' different levels of analysis. Those listed above are only the most apparent. Neither theory is better or worse than the other. Overall, they are so much alike that they beg for consolidation. It will be interesting to see if that happens.

**Acknowledgement.** None.

**Financial support.** This work received no specific grant from any funding agency, commercial or not-for-profit sectors.

**Competing interest.** None.

## References

Atkinson, R. F. (1978). *Knowledge and explanation in history: An introduction to the philosophy of history*. Cornell University Press.
Beach, L. R. (2009). Decision making: Linking narratives and action. *Narrative Inquiry*, *19*(2), 293–416.
Beach, L. R. (2010). *The psychology of narrative thought*. Xlibris.
Beach, L. R. (2018). Narrative thought and management. *Organizational Dynamics*, *47*, 63–69.
Beach, L. R. (2019). *The structure of conscious experience*. Cambridge Scholars.
Beach, L. R., Bissell, B. L., & Wise, J. A. (2016). *A new theory of mind*. Cambridge Scholars.
Beach, L. R., & Wise, J. A. (2022). *The theory of narrative thought*. Cambridge Scholars.
Carroll, N. (2001). On the narrative connection. In W. van Peer & S. Chatman (Eds.), *New perspectives on narrative perspective*. State University of New York Press.
Phillips, L. D. (1970). The true probability problem. *Acta Psychologica*, *34*, 254–264.

# High-stakes decisions do not require narrative conviction but narrative flexibility

Fritz Breithaupt[a,b], Milo Hicks[b,c], Benjamin Hiskes[a,b,c] and Victoria Lagrange[b,d]

[a]Cognitive Science Program, Indiana University, Bloomington, IN 47405, USA; [b]Experimental Humanities Lab, Indiana University, Bloomington, IN 47405, USA; [c]Department of English, Indiana University, Bloomington, IN 47405, USA and [d]Department of English, Kennesaw State University, Kennesaw, GA 30144, USA.
fbreitha@indiana.edu
milohick@iu.edu
bhiskes@iu.edu
vlagrang@kennesaw.edu
www.experimentalhumanities.com

doi:10.1017/S0140525X22002606, e85

**Abstract**

We challenge Johnson et al.'s assumption that people reduce unclear situations to a single narrative explanation and that such reduction would be adaptive for decision-making under radical uncertainty. Instead, we argue that people imagine and maintain multiple narrative possibilities throughout the decision-making process and that this process provides cognitive flexibility and adaptive benefits within the proposed model.

The target article by Johnson et al. provides a framework for decision-making under radical uncertainty that relies on narratives to create the "single likeliest" chain of events (target article, sect. 3, para. 11). However, it does not leverage a central feature of narrative thinking: Our ability to imagine multiple possible pasts and futures. While Conviction Narrative Theory (CNT) posits the simulation of more than one imagined future, it also assumes that explanatory reasoning results in a single narrative that is "retained for downstream computations such as forecasting" (target article, sect. 7.2, para. 2) and that simulation generates "only one future for each choice" (target article, sect. 3, para. 7). We challenge the assumption that people reduce unclear situations to a single narrative explanation and that such reduction would be adaptive for decision-making. In other words, we disagree that "stories resist Heisenberg's principle – they take only one state at a time" (target article, sect. 7.2, para. 5). We contend that the plurality of narrative thinking should be integrated more fully into the various processes of CNT: Not only do we simulate more than one imagined future for each choice, but we also make sense of the present by relying on multiple narratives at the explanatory stage of decision-making.

Let us consider the authors' example of the widow hearing a noise inside her home, with a slight modification, to illustrate the importance of multiple narratives in decision-making. The widow hears a noise coming from downstairs late at night and considers two competing explanations: The noise could be a burglar breaking in, or it could be her lover entering unannounced, but she cannot know for certain. Even if she adopts whichever narrative "feels right," the alternative explanation remains in the back of her mind. She then simulates several possible courses of action: She wants to defend herself; however, if it is her lover, she does not want to attack them. If the widow follows CNT, she would make a choice and bet on "the single likeliest imagined future" (target article, sect. 3, para. 11). Either she would grab and swing her baseball bat or she would jump forward to hug her lover. Obviously, a false choice here would be dangerous.

This example highlights that in many situations, it may be necessary to simultaneously consider not only multiple simulations of the future, but also multiple explanations of the present. If narratives play a key role in decision-making, then to understand this process we should consider how people make sense of them. How people consume narratives does not necessarily indicate how people produce them to make decisions. However, the process of narrative consumption provides insight into the relevance of narratives for decision-making because receiving a narrative also involves interpreting an ongoing situation and making predictions for the narrative's future (Campion, 2004; Magliano, Dijkstra, & Zwaan, 1996).

In research on narrative processing, there is in fact evidence that people *do* consider multiple possible versions of narratives simultaneously. Every narrative involves expectation management for the future, and thereby includes at least a minimal degree of suspense. Suspense involves simultaneously and continuously imagining at least two possibilities for a narrative's future, namely a feared outcome and a desired outcome (Carroll, 1990). Intriguingly, suspense and narrative enjoyment remain largely intact even when the outcome of a narrative is known (Carroll, 2001; Johnson & Rosenbaum, 2015; Leavitt & Christenfeld, 2011). One potential explanation for this paradox is that the narrative consumption experience is not only about accurately predicting a single outcome of the story, but also imagining alternative possibilities (Hiskes, Hicks, Evola, Kincaid, & Breithaupt, 2022). This view of narrative processing as possibility generation helps explain the popularity of fan fictions that aim at playing-out alternative versions of a story, as well as economic scenarios involving rapidly shifting opinions about Bitcoin, to use Robert J. Shiller's example (2019), since competing narratives are already available.

Just as narrative processing research indicates that people may hold multiple future narratives in mind simultaneously, there is also evidence that people maintain possible alternatives for a narrative's past using counterfactuals. De Vega, Urrutia, and Rifo (2007) find that immediately after reading, counterfactual and factual information remains separate but accessible in memory to readers. Similarly, Ferguson and Jayes (2018) find that readers do indeed attend to and evaluate counterfactual information, even when it differs from factual narrative situations. Consequently, there is evidence that when consuming narratives, we do not remain "blind to alternative possibilities" as Johnson et al. contend (target article, sect. 7.2, para. 1). Instead, alternative explanations of a situation remain open and available.

While this research addresses actual narrative processing, we argue that considering multiple possibilities in real-life decisions is also adaptively beneficial. The authors insist on the value of conviction in decision-making, while highlighting the subsequent need for learning and reassessment based on new information. However, conviction may be detrimental in quickly changing situations, and constant reappraisal and reassessment may be costly as well. If a situation is radically uncertain, settling on a single narrative explanation may pose a high risk, as the example with the widow demonstrates. Instead, keeping multiple possible narratives in mind minimizes the chances of being surprised by, and thus unprepared for, changing circumstances. Moreover, such narrative multiplicity offers a way to compare and discriminate between different courses of action at both the evaluation and simulation stage. In other words, keeping multiple narratives open in decision-making lowers the risks posed by overcommitting to a single narrative that may ultimately offer bad courses of action.

Thus, we recommend a modification to CNT: Considering multiple narrative possibilities throughout the decision-making process is not only something that people do, but it is also beneficial for evaluating the best possible decision. This multiplicity does not preclude decision-making but allows for flexibility as we try to understand our present through explanation, imagine our future through simulation, as well as make decisions through affective evaluation.

**Author's contribution.** All authors contributed equally.

**Financial support.** This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

**Competing interest.** None.

## References

Campion, N. (2004). Predictive inferences are represented as hypothetical facts. *Journal of Memory and Language*, *50*(2), 149–164. https://doi.org/10.1016/j.jml.2003.10.002
Carroll, N. (1990). *The philosophy of horror: Or, paradoxes of the heart.* Routledge.
Carroll, N. (2001). The paradox of suspense. In *Beyond aesthetics: Philosophical essays* (pp. 254–270). Cambridge University Press. https://doi.org/10.1017/CBO9780511605970.017
De Vega, M., Urrutia, M., & Rifo, B. (2007). Canceling updating in the comprehension of counterfactuals embedded in narratives. *Memory & Cognition*, *35*(6), 1410–1421. https://doi.org/10.3758/BF03193611
Ferguson, H. J., & Jayes, L. T. (2018). Plausibility and perspective influence the processing of counterfactual narratives. *Discourse Processes*, *55*(2), 166–186. https://doi.org/10.1080/0163853X.2017.1330032

Hiskes, B., Hicks, M., Evola, S., Kincaid, C., & Breithaupt, F. (2022). Multiversionality: Considering multiple possibilities in the processing of narratives. *Review of Philosophy and Psychology*. https://doi.org/10.1007/s13164-022-00621-5

Johnson, B. K., & Rosenbaum, J. E. (2015). Spoiler alert: Consequences of narrative spoilers for dimensions of enjoyment, appreciation, and transportation. *Communication Research*, *42*(8), 1068–1088. https://doi.org/10.1177/0093650214564051

Leavitt, J. D., & Christenfeld, N. J. (2011). Story spoilers don't spoil stories. *Psychological Science*, *22*(9), 1152–1154. https://doi.org/10.1177/0956797611417007

Magliano, J. P., Dijkstra, K., & Zwaan, R. A. (1996). Generating predictive inferences while viewing a movie. *Discourse Processes 22*(3): 199–224. https://doi.org/10.1080/01638539609544973

Shiller, R. J. (2019). *Narrative economics: How stories go viral and drive major economic events*. Princeton University Press.

# Conviction Narrative Theory gains from a richer formal model

Leigh Caldwell

First Floor, Sackville House, 143-149 Fenchurch St, London EC3M 6BL, UK.
leigh@irrationalagency.com
www.leighcaldwell.com

doi:10.1017/S0140525X2200262X, e86

**Abstract**

Conviction Narrative Theory (CNT) is a convincing descriptive theory, and Johnson et al.'s formal model is a welcome contribution to building more precise, testable hypotheses. However, some extensions to the proposed model would make it better defined and more powerful. The suggested extensions enable the model to go beyond CNT, predicting choice outcomes and explaining affective phenomena.

Narratives have long been recognised as important by psychologists (De Beaugrande & Colby, 1979; Bruner, 1985; Pennington & Hastie, 1986; Sarbin, 1986) and more recently by economists (Eliaz & Spiegler, 2020; Shiller, 2017). Intuitively and empirically, narrative plays an important role in thinking.

Conviction Narrative Theory (CNT) (target article; Tuckett and Nikolic, 2017) convincingly proposes that people use narratives to settle on a course of action in the face of uncertainty. Earlier work on CNT has not clearly defined what a narrative is: The target article's significant new contribution is a formal model of narrative. Its model of a narrative can be summarised as a graph of objects, in which:

- Each object may have positive, negative or neutral valence,
- Any pair may be causally related,
- Any pair may be temporally related,
- Any subgraph can represent an analogy with a second, isomorphic subgraph.

This model helps to formalise the phenomena previously treated only descriptively. As it stands, however, it leaves some matters unclear. The model is not quite full enough to support some of the phenomena described in the target article.

The authors ask, in the context of probabilistic utility theory, *where do these numbers come from?* One might ask them in turn: *Where do these narratives come from?* The answer, that they are "supplied in part by the social environment" is imprecise and incomplete.

Narratives are to be used for affective evaluation – but other than an optional binary valence on each node, there is no affective component in the model. During explanation, agents evaluate whether a narrative "feels right," but the model contains nothing that they can use to make this judgement – this feeling must be imported from outside.

This missing information renders some of the examples ambiguous. In Figure 6c, although Fundamentals have a causal effect on Price, this could be either positive or negative – the diagram cannot capture this. In others, the assignment of valence is not explained. In Figure 5a, Masks receive positive valence. Is this because people like wearing them? Or is the valence inferred from Masks' negative effect on infection? The latter would require narratives to be dynamic rather than static.

Three extensions to the model would resolve these issues. The mathematical approach is provided by the "fuzzy graph" literature (Blue, Bush, & Puckett, 2002; Kóczy, 1992; Yeh & Bang, 1975; Zadeh, 1999). A fuzzy graph's nodes and edges take non-binary values, usually a real number in [0, 1].

First, where the models come from: Instead of constructing narratives on-the-fly for each decision, a persistent causal graph is specified, representing the agent's whole mental model of the world. This graph is learned (by conditioning, or through verbal learning). Narratives are selected as subgraphs of this model (perhaps with minor modifications), not created from scratch.

The authors relate the story of Mrs O'Leary's cow, who kicked over a lantern that started the Great Chicago Fire. Although this narrative is "given" to us in the telling, it is believable because we already know that cows kick lanterns and lanterns start fires. Without prior causal beliefs, it might invite doubt rather than conviction.

Narrative subgraphs of a pre-existing causal graph can be instantiated and used faster, and more easily communicated to other people, than new narratives. This explains why a chosen narrative is more likely to contain concepts already salient to the agent (Boyer, 2003). The authors hint at this: "…the most contagious narratives including a small number of novel concepts against a larger background of familiar concepts…in digestible form because they match our default causal-temporal representations of events…"

Second, to enable affective evaluation, valence in the graph is replaced with a scalar affective value, represented by degree of shading in Figures 1 and 2 (interpreted as "anticipated reward" as in Berns, Laibson, & Loewenstein, 2007). When a narrative is evaluated during mental simulation, this affective value is experienced (Ainslie, 2017) and that experience is used in evaluation.

Each affective value is updated during mental simulation as its downstream causal consequences are experienced, allowing future simulations to be more efficient and accurate (Schultz, Dayan, & Montague, 1997). Affectively appealing narratives become "attractors" (Parunak, 2022).

Third, to better specify causal influence, the edges are assigned a scalar coefficient reflecting strength or likelihood of causality. The authors rightly observe that exact probabilities are incalculable under radical uncertainty, but certainly agents can understand that some causal links are stronger than others (Sloman & Lagnado, 2015). These coefficients can influence how narratives are mentally simulated, even without conscious awareness of the agent.

This change allows narratives to be traversed by more than one possible pathway. Imagine a brain evaluating multiple pathways in parallel (or by quickly switching between them)

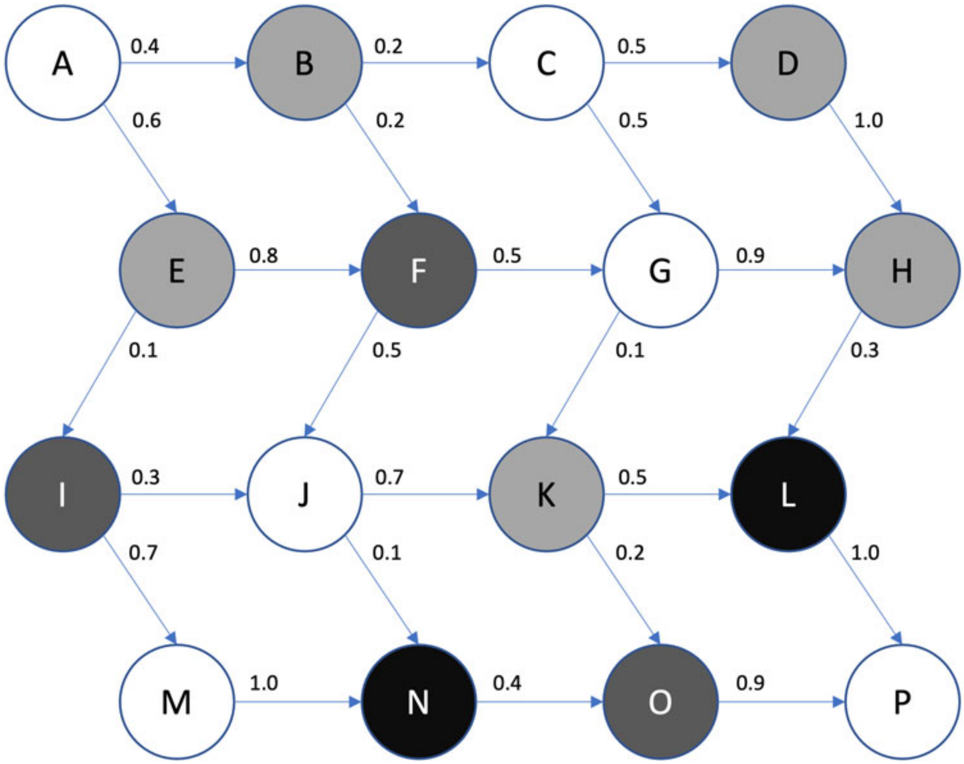


**Figure 1** (Caldwell). The background graph prior to narrative selection.

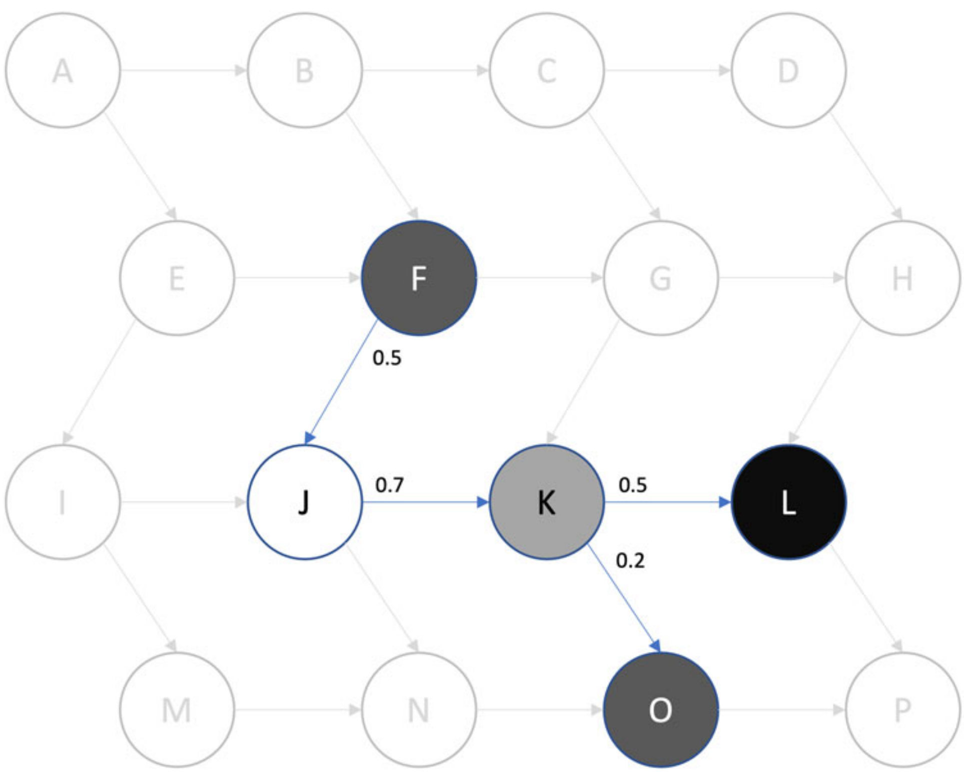


**Figure 2** (Caldwell). A narrative selected as subgraph, prompted by a choice situation.

rather than in isolation. CNT's proposed explanation and simulation phases then merge into one process, in which multiple narrative pathways are tried on for both plausibility and affective appeal, until one preferred option is found. This perhaps better reflects the human experience of struggling with uncertainty.

These refinements more fully implement CNT, and enable the model to predict more than just conviction. In offline replay

https://doi.org/10.1017/S0140525X22001157 Published online by Cambridge University Press

(Momennejad, Otto, Daw, & Norman, 2018), by contrasting multiple narratives of present and future, agents can update their mental models even when not choosing between options. If mental simulation over this narrative model generates "synthetic reward" (Caldwell, 2018a, 2018b), daydreaming (Schelling, 1987), enjoying memory replay and transportation by fiction (Polichak & Gerrig, 2002) can all be explained. The How-Does-It-Feel heuristic (Caldwell, 2018b; Pham, 1998) is implemented by this model.

In commercial work, a similar formal model ("System 3") has been used (Caldwell & Seear, 2019) to predict frequency of shopping behaviour, responses to advertising messages and consumer sustainability behaviours.

In their conclusion, the authors anticipate two critiques: That their theory may be seen as too grandiose, or too skeletal. This commentary lands on the "skeletal" side, but perhaps this proposed extension of the formal model can add some meat to those very promising bones.

**Financial support.** This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

**Competing interest.** None.

## References

Ainslie, G. (2017). De Gustibus Disputare: Hyperbolic delay discounting integrates five approaches to impulsive choice. *Journal of Economic Methodology*, *24*(2), 166–189.
Berns, G. S., Laibson, D., & Loewenstein, G. (2007). Intertemporal choice – toward an integrative framework. *Trends in Cognitive Sciences*, *11*(11), 482–488.
Blue, M., Bush, B., & Puckett, J. (2002). Unified approach to fuzzy graph problems. *Fuzzy Sets and Systems*, *125*(3), 355–368.
Boyer, P. (2003). Religious thought and behaviour as by-products of brain function. *Trends in Cognitive Sciences*, *7*(3), 119–124.
Bruner, J. (1985). Chapter VI: Narrative and paradigmatic modes of thought. *Teachers College Record*, *86*(6), 97–115.
Caldwell, L. (2018a). *Reinforcement learning made me irrational* [Conference presentation]. TIBER 2018, Tilburg, Netherlands.
Caldwell, L. (2018b). *Inferring heuristics from complex causal models* [Conference presentation]. International Conference: Heuristics in Organisations and Society, Herbert Simon Society, Turin, Italy.
Caldwell, L., & Seear, L. (2019). System 3: Measuring the consumer's imagination. ESOMAR Congress 2019.
De Beaugrande, R., & Colby, B. N. (1979). Narrative models of action and interaction. *Cognitive Science*, *3*(1), 43–66.
Eliaz, K., & Spiegler, R. (2020). A model of competing narratives. *American Economic Review*, *110*(12), 3786–3816.
Kóczy, L. (1992). Fuzzy graphs in the evaluation and optimization of networks. *Fuzzy Sets and Systems*, *46*(3), 307–319.
Momennejad, I., Otto, A. R., Daw, N. D., & Norman, K. A. (2018). Offline replay supports planning in human reinforcement learning. *elife*, *7*, e32548.
Parunak, H. V. D. (2022). Model mechanisms and behavioral attractors. In *Social Simulation Conference (SSC2022), pages (under review), University of Milan, Milan, Italy.*
Pennington, N., & Hastie, R. (1986). Evidence evaluation in complex decision making. *Journal of Personality and Social Psychology*, *51*(2), 242.
Pham, M. T. (1998). Representativeness, relevance, and the use of feelings in decision making. *Journal of Consumer Research*, *25*(2), 144–159.
Polichak, J. W., & Gerrig, R. J. (2002). Get up and win! Participatory responses to narratives. In M. C. Green, J. J. Strange, & T. C. Brock (Eds.), *Narrative impact: Social and cognitive foundations* (pp. 71–95). Erlbaum.
Sarbin, T. R. (1986). *Narrative psychology: The storied nature of human conduct*. Praeger Publishers/Greenwood Publishing Group.
Schelling, T. C. (1987). The mind as a consuming organ. In J. Elster (Ed.), *The Multiple Self* (pp. 177–196). Cambridge University Press.
Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science*, *275*(5306), 1593–1599.
Shiller, R. J. (2017). Narrative economics. *American Economic Review*, *107*(4), 967–1004.
Sloman, S. A., & Lagnado, D. (2015). Causality in thought. *Annual Review of Psychology*, 66, 223–247.
Tuckett, D., & Nikolic, M. (2017). The role of conviction and narrative in decision-making under radical uncertainty. *Theory & Psychology*, *27*(4), 501–523.
Yeh, R. T., & Bang, S. Y. (1975). Fuzzy relations, fuzzy graphs, and their applications to clustering analysis. In *Fuzzy sets and their applications to cognitive and decision processes* (pp. 125–149). Academic Press.
Zadeh, L. A. (1999). Fuzzy logic and the calculi of fuzzy rules, fuzzy graphs, and fuzzy probabilities. *Computers & Mathematics with Applications*, *37*(11–12), 35.

# Epistemic trust and unchanging personal narratives

Chloe Campbell[a] and Peter Fonagy[b]

[a]Research Department of Clinical, Educational and Health Psychology, University College London, London WC1E 7HB, UK and [b]Division of Psychology and Language Sciences, University College London, London WC1H 0AP, UK.
c.campbell@ucl.ac.uk
p.fonagy@ucl.ac.uk
https://www.ucl.ac.uk/psychoanalysis/people/chloe-campbell
https://www.ucl.ac.uk/pals/people/peter-fonagy

doi:10.1017/S0140525X22002540, e87

**Abstract**

Focusing on imagination and the social context in the generation of conviction narratives, we propose that these elements are dynamically related to one another, and crucially that it is the nature of this relationship that determines individuals' level of epistemic openness and capacity to respond adaptively to update their narratives in a way that increases the possibility of more successful decision-making.

Johnson et al. have taken a radical step in relation to "radical uncertainty." They have created a model for understanding the wide body of research demonstrating that individuals tend to overestimate their own rationality and decision-making capacities, including, according to the authors, academics, politicians and economists, who collectively tend to have overestimated human rationality. Their case is profound in arguing that the network of norms and institutions society gathers into a system of rules for identifying truth – what Jonathan Rauch, a senior fellow at the Brookings Institution, recently termed *a constitution of knowledge* (Rauch, 2018) – is mistaken in relation to understanding how decisions are made.

Here we seek to extend the implications of Conviction Narrative Theory in relation to the role of imagination and its relationship with the social context. Johnson et al. argue that both factors are insufficiently incorporated into decision-making theory: We additionally propose that they are dynamically related to one another, and that it is the nature of this relationship that determines individuals' capacity to respond adaptively to new information. One defining difficulty with conviction narratives is that, notwithstanding evidence that challenges their position, individuals are often unable to adjust their narratives. One way of understanding this is irrationality; another way is through petrification of the narrative in the light of epistemic mistrust: New information that challenges the narrative is understood but not adopted into what Polanyi (1959/2014) calls "personal knowledge," the sum total of what any individual knows, because it is not regarded as relevant (Fonagy, Luyten, & Allison, 2015). Inflexible, dysfunctional conviction narratives are indications of a lack of trust in incompatible but pertinent information. We understand the possible reasons for this epistemic closure and unwarranted certainty as inter-relational in nature (Fonagy et al., 2022). In a developmental psychopathology context, we described the factors that drive this entrenchment as early relational insecurity (Fonagy, Luyten, Allison, & Campbell, 2017a, 2017b). But we have also more broadly come to understand this epistemic shut-down as an outcome of functioning within a social environment in which the individual does not feel recognised in the communications they receive. We have suggested that such lack of experience of recognition in others' narratives in relation to the self may constitute something of the "missing link" in understanding the relationship between socio-economic deprivation and/or alienation and psychopathology (Fonagy et al., 2022).

Epistemic dysfunction does not exclusively manifest as outright mistrust, it also presents as lack of discrimination in relation to the communication of knowledge, rendering the individual vulnerable to exploitation or misinformation, which we have termed epistemic credulity (Campbell et al., 2021). In a state of epistemic credulity, individuals misread or are misled by the narratives presented by others, imagining that they are being recognised and understood, and thus reducing adaptive epistemic vigilance towards the information they receive (Sperber et al., 2010). We might think here of a populist politician who, despite their blatant self-interest, creates a powerful narrative of being able to represent the sense of betrayal and injustice felt by the ordinary person, to the extent that obvious misinformation about election results is believed.

It is easy to pathologise these processes as they manifest in their most egregious forms, perhaps partially driven by a social media-escalated breakdown in epistemic consensus. In fact, such ways of thinking together hold important functions in maintaining social stability (Mercier & Sperber, 2017) in an expectable environment (Cicchetti & Lynch, 1995). Community and collaboration are supported by a set of mental processes reserved for shared cognition or relational mentalising, also described as the "we-mode." When we intend to accomplish an outcome jointly with others, we adopt a "first-person plural perspective" – the we-mode (Gallotti & Frith, 2013, p. 160). The we-mode may be organised around cognitive and neural structures that are intrinsic to individual make-up and are the product of a distinct developmental and evolutionary history (Tomasello, 2019). Building on joint intentionality, the *joint agent* emerges where mental states are aligned to achieve a common goal, grounded in respect born of each having a role in the collaborative activity (Tomasello, 2016). The route by which the we-mode is triggered may be understood thus: (1) The learner's *imagined* sense of self (their personal narrative) (2) is imagined by the instructor, establishing a prospect for the we-mode and (3) this image is perceived by the learner, reinforcing the potential we-mode and (4) compared with the learner's personal narrative and (5) in case of a match, co-representation is created: The we-mode removes the I-mode's protection from change and the channel for rapid, efficient modification of personal knowledge is opened. Relational mentalising is thus key to establishing epistemic trust, it is dependent on both instructor and learner taking an imaginative leap in relation to the other's mind, but using that imagination in a manner that is contained within a reasonably accurate assessment of both one's own and the other's mental state. Mistrust might be understood as a failure of imagination, credulity as manipulated or untethered imagination.

Johnson et al. describe narratives as representations in individual minds; shared narratives as being formed by the common elements fragmentarily shared across a social group. We suggest that it is harder to separate individual from shared narratives and knowledge than this conceptualisation suggests. Conviction narratives are particularly relevant for social contexts where strong boundaries around social groups exist and information incompatible with current beliefs are discarded if it comes from individuals who are members of an out-group to whom epistemic trust is not extended (Tong, Wang, & Danovitch, 2020). As information incompatible with beliefs accumulates, paradoxically mistrust increases as beliefs come to serve to identify the social group who hold them. This is an outcome of epistemic vigilance (sometimes hypervigilance) precluding the integration of socially conveyed information into knowledge schemata. We have recently argued that models for understanding social processes and social cognition have failed to learn from developmental psychopathology (Campbell & Fonagy, 2022); here we perhaps make a bolder point that the psychopathology of everyday life – the lapses in communication, ruptures and the outcomes of inhibited or disorderly imagination – is the engine of culture, because these processes create the narratives that hold (or fail to hold) groups together.

**Financial support.** PF is in part supported by the NIHR Applied Research Collaboration (ARC) North Thames at Barts Health NHS Trust. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

**Competing interest.** None.

## References

Campbell, C., & Fonagy, P. (2022). Bifocalism is in the eye of the beholder: Social learning as a developmental response to the accuracy of others' mentalizing. *Behavioral and Brain Sciences*, *45*, E254. https://doi.org/10.1017/S0140525X22001297

Campbell, C., Tanzer, M., Saunders, R., Booker, T., Allison, E., Li, E., … Fonagy, P. (2021). Development and validation of a self-report measure of epistemic trust. *PLOS ONE*, *16*(4), e0250264. https://doi.org/10.1371/journal.pone.0250264

Cicchetti, D., & Lynch, M. (1995). Failures in the expectable environment and their impact on individual development: The case of child maltreatment. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology* (Vol. *2*, pp. 32–71). Wiley.

Fonagy, P., Campbell, C., Constantinou, M., Higgitt, A., Allison, E., & Luyten, P. (2022). Culture and psychopathology: An attempt at reconsidering the role of social learning. *Development and Psychopathology*, *34*(4), 1205–1220. https://doi.org/10.1017/S0954579421000092

Fonagy, P., Luyten, P., & Allison, E. (2015). Epistemic petrification and the restoration of epistemic trust: A new conceptualization of borderline personality disorder and its psychosocial treatment. *Journal of Personality Disorders*, *29*(5), 575–609. https://doi.org/10.1521/pedi.2015.29.5.575

Fonagy, P., Luyten, P., Allison, E., & Campbell, C. (2017a). What we have changed our minds about: Part 1. Borderline personality disorder as a limitation of resilience. *Borderline Personality Disorder and Emotion Dysregulation*, *4*, 11. https://doi.org/10.1186/s40479-017-0061-9

Fonagy, P., Luyten, P., Allison, E., & Campbell, C. (2017b). What we have changed our minds about: Part 2. Borderline personality disorder, epistemic trust and the developmental significance of social communication. *Borderline Personality Disorder and Emotion Dysregulation*, *4*, 9. https://doi.org/10.1186/s40479-017-0062-8

Gallotti, M., & Frith, C. D. (2013). Social cognition in the we-mode. *Trends in Cognitive Sciences*, *17*(4), 160–165. https://doi.org/10.1016/j.tics.2013.02.002

Mercier, H., & Sperber, D. (2017). *The enigma of reason: A new theory of human understanding*. Allen Lane.

Polanyi, M. (1959/2014). *The study of man*. University of Chicago Press.

Rauch, J. (2018). The constitution of knowledge. *National Affairs*, *37*(Fall), 125–137.

Sperber, D., Clement, F., Heintz, C., Mascaro, O., Mercier, H., Origgi, G., & Wilson, D. (2010). Epistemic vigilance. *Mind & Language*, *25*(4), 359–393. https://doi.org/10.1111/j.1468-0017.2010.01394.x

Tomasello, M. (2016). *A natural history of human morality*. Harvard University Press.

Tomasello, M. (2019). *Becoming human: A theory of ontogeny*. The Belknap Press of Harvard University Press.

Tong, Y., Wang, F., & Danovitch, J. (2020). The role of epistemic and social characteristics in children's selective trust: Three meta-analyses. *Developmental Science*, *23*(2), e12895. https://doi.org/10.1111/desc.12895

# Embodied choices bypass narratives under radical uncertainty

Rouwen Cañal-Bruland[a] and Markus Raab[b,c]

[a]Department for the Psychology of Human Movement and Sport, Friedrich Schiller University Jena, 07749 Jena, Germany; [b]Department of Performance Psychology, German Sport University Cologne, 50933 Cologne, Germany and [c]School of Applied Sciences, London South Bank University, SE1 0AA London, UK.
rouwen.canal.bruland@uni-jena.de
raab@dshs-koeln.de
https://www.spowi.uni-jena.de/en/departments/department-for-the-psychology-of-human-movement-and-sport/people/rouwen-canal-bruland
https://fis.dshs-koeln.de/portal/en/persons/markus-raab(05e809e4-1be0-45aa-b811-58adc6779693).html

doi:10.1017/S0140525X22002722, e88

**Abstract**

Johnson et al. suggest that we rely on narratives to make choices under radical uncertainty. We argue that in its current version Conviction Narrative Theory (CNT) does not account for embodied, direct sensorimotor influences on choices under radical uncertainty that may bypass narratives, particularly in highly time-constrained situations. We therefore suggest to extend CNT by an embodied choice perspective.

Johnson et al. suggest that Conviction Narrative Theory (CNT) explains choices under radical uncertainty, arguing that we rely on *narratives* to make decisions. While we agree with many aspects of CNT and claims made in this target article, including the suggested relationship with boundedly rational approaches, we argue that CNT needs to be extended to account for direct sensorimotor influences and hence embodied choices that generate and inform decisions. More specifically, we challenge the proposition that "narratives characterize real-world decisions under radical uncertainty" (target article, Table 2). In our view, this proposition is not generalizable because narratives do not include embodied choices defined as choices in which the sensorimotor system itself generates and informs decisions (Raab, 2017). Embracing an embodied choices perspective may become particularly relevant if CNT also aims at explaining and predicting motor choices and decisions that are made on rather short time scales.

To start with, narratives are defined as "structured, higher-order mental representations" that serve to "explain the past, […] predict the future, and evaluate possible futures." We do not dispute that CNT may account for a plethora of real-world decisions under radical uncertainty, such as for the chosen empirical examples regarding macro- and microeconomic decision-making. However, CNT relies on the assumption that narratives serve to first make decisions and then effectuate those decisions by means of actions. In other words, CNT can be classified as a classical "first decide, then act"-theory (Gordon et al., 2021; Wispinski, Gallivan, & Chapman, 2020). Again, this classical approach may be valid for decision-making in economic decision-making under radical uncertainty, yet it falls short to explain many choices "our brains evolved to deal with," namely "embodied decisions" (Gordon et al., 2021, p. 722).

First, embodied decisions include many decisions in daily life that are made during continuous movements that are characterized by constantly changing action dynamics. Such changes in bodily action dynamics have been shown to directly influence decision-making (Grießbach, Incagli, Herbort, & Cañal-Bruland, 2021). It seems questionable whether such effects can be explained by narratives as the representational "currency of thought." For instance, Jax and Rosenbaum (2007) found uneconomic (or perhaps cognitively irrational) motor behaviors in reaching tasks. They explained their findings by arguing that participants faced the problem to trade two incommensurate currencies, namely motor programming and planning costs (and associated cognitive and neural resources) and biomechanical costs, referred to as "apples and oranges." It follows that by not accounting for the influence of sensorimotor information and embodied choices, CNT may fall short to explain at least some real-world decisions under radical uncertainty.

Second, data from the external world need to be perceived in order to inform narratives. Similar to embodied decisions there is evidence to suggest that perception is embodied (Lepora & Pezzulo, 2015). For example, Proffitt (2006) reviewed evidence corroborating that the perception of the environment changes depending on anticipated action costs, thereby adding to our critique that CNT is lacking an embodied perspective to explain choices under radical uncertainty.

Third, next to the emerging fields of embodied decisions and embodied perception, there is growing awareness that bounded rationality and heuristics are embodied (Gigerenzer, 2021). For instance, motor heuristics have been defined as simple rules of thumb that allow actors to motorically choose between options that satisfy the current demands of a given task or situation (Raab, 2017). In contrast to the examples chosen to support CNT, however, whenever movements are part of the action response, it is not only important what to choose but how the body constraints or gives choices when exact movements in time and space matter such as when driving a car, crossing a street, or playing sports.

Together, research on embodied decision, embodied perception as well as embodied choices points to the need to extend CNT in order to account for the embodied nature of choices under radical uncertainty. In our view, this is likely to be particularly important for explaining choices in highly time-constrained situations such as in fastball sports. Reflecting on the representations and processes in CNT illustrated in Figure 2, we are skeptical that such elaborate, sequential steps are run through, for instance, when a batter chooses when and where to move the bat in order to hit the ball into the outfield. We argue that it is much more likely that such choices bypass narratives to allow for successful movement solutions under high temporal demands.

In concluding their target article, Johnson et al. raise the possibility that there may be other forms of information they may not have considered. Here, we suggested that CNT would benefit from an embodied choice perspective that does not take the body and motor actions as the mere executor of cognitive decisions, but acknowledges that changing action dynamics are generating and informing embodied choices. If such an embodied choice perspective would be embraced, CNT may also account for highly time-constrained situations which by definition reflect choice under radical uncertainty, thereby allowing to generalize CNT to a broader set of real-world decisions.

**Financial support.** This work was supported by the German Research Foundation (DFG), (RCB, CA 635/2-2, CA 635/4-1), (MR, RA 940/15-2).

**Competing interest.** None.

## References

Gigerenzer, G. (2021). Embodied heuristics. *Frontiers in Psychology*, *12*, 711289.
Gordon, J., Maselli, A., Lancia, G. L., Thiery, T., Cisek, P., & Pezzulo, G. (2021). The road towards understanding embodied decisions. *Neuroscience & Biobehavioral Reviews*, *131*, 722–736.
Grießbach, E., Incagli, F., Herbort, O., & Cañal-Bruland, R. (2021). Body dynamics of gait affect value-based decisions. *Scientific Reports*, *11*, 11894.
Jax, S. A., & Rosenbaum, D. A. (2007). Hand path priming in manual obstacle avoidance: Evidence that the dorsal stream does not only control visually guided actions in real time. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 425–441.
Lepora, N. F., & Pezzulo, G. (2015). Embodied choice: How action influences perceptual decision making. *PLoS Computational Biology*, *11*, e1004110.
Proffitt, D. R. (2006). Embodied perception and the economy of action. *Perspectives on Psychological Science*, *1*, 110–122.
Raab, M. (2017). Motor heuristics and embodied choices: How to choose and act. *Current Opinion in Psychology*, *16*, 34–37.
Wispinski, N. J., Gallivan, J. P., & Chapman, C. S. (2020). Models, movements, and minds: Bridging the gap between decision making and action. *Annals of the New York Academy of Sciences*, *1464*(1), 30–51.

# The ephemeral stories of our lives

Nick Chater

Behavioural Science Group, Warwick Business School, University of Warwick, Coventry CV4 7AL, UK.
nick.chater@wbs.ac.uk

doi:10.1017/S0140525X22002552, e89

**Abstract**

Johnson et al. make a persuasive case that qualitative, story-like reasoning plays a crucial role in everyday thought and decision-making. This commentary questions the cohesiveness of this type of reasoning and the representations that generate it. Perhaps narratives do not underpin, but are ephemeral *products* of thought, created when we need to justify our actions to ourselves and others.

Johnson et al. make a compelling case that what they call conviction narratives play an important role in human thought and decision-making. But are such narratives primarily *products* of more basic cognitive representations and processes, which are constructed in the moment to justify and persuade?

Humans continually tell stories about themselves and others, whether in conversation, courts of law, or literature. But, of course, they generate not merely stories, but external representations of many kinds: Diagrams, maps, sketches, descriptions, theories, hypotheses, regulations, algorithms, mathematical proofs, logical formalisms, and many more.

A central and classic question for the behavioral and brain sciences is which of these representations, if any, provide useful analogs of the internal representations underlying thought. Do animals, including humans, represent spatial layouts using "mental maps" (Tolman, 1948)? Is mental imagery underpinned by a distinctively pictorial style of representation (Kosslyn, 1980; Pylyshyn, 1981)? Does the generation and understanding of natural language involve the translation to and from an internal language of thought (Dennett, 1978; Fodor, 1975)? Is our naïve understanding of other people, or the external world, rooted in internally represented theories of folk psychology or naïve physics (Bloom, 2005)? Similarly, researchers have long wondered whether aspects of everyday knowledge are organized into standardized, perhaps story-like, scripts (e.g., Schank & Abelson, 1977) – a viewpoint that Johnson et al. explore in great depth.

How can we tell? As the classic debates mentioned above suggest, there is no single accepted and definitive approach. But one direct approach is to ask what it is, and what it is not, possible to represent, using different types of representation. Suppose that a person judges London to be west of Tokyo, Tokyo to be west of San Francisco, and San Francisco to be west of London. This immediately implies that these judgments cannot be being directly read off from a flat 2D map-like internal representation of the geography of the world – because it is impossible to represent these relationships on a 2D map. Of course, those judgments might be read off from an internal representation that, like the earth itself, has the form of a sphere; or they might be stored in a purely symbolic database.

Consider an ingenious study by Muryy and Glennerster (2021), in which people navigate a virtual reality environment, in which "wormholes" have been created which are inconsistent 3D space. If people represent the local environment using Euclidean spatial maps, then they should be unable to represent such wormholes, should therefore be unable to represent such environments using their mental maps,with the result that their navigation should be heavily impeded. In practice, though, people are able to move around these strange spaces successfully, apparently oblivious to their contradictions with the assumptions of Euclidean space. Or consider the visual representation of so-called "impossible objects," such as the Penrose triangle (Penrose & Penrose, 1958). If vision, or mental imagery, generally involves conjuring up an iconic pictorial representation in 3D space, then such a process should be confounded by impossible objects, because no consistent representation can be found. In practice, the visual system is only able to detect such inconsistencies only after considerable scrutiny.

Turning to decision-making, the assumption that outcomes are assigned cardinal-valued utilities similarly implies that intransitive preferences (preferring A to B, B to C, and C to A) should be impossible to represent, and should be observed empirically due only to unstable preferences or noise. But many have argued that choices can often be systematically intransitive, and hence cannot be mediated by any kind of utility representation over possible outcomes (e.g., Tsetsos et al., 2016).

In each of these cases, the choice of basic representational units determines what can, and what cannot, be represented – and in each of these cases, observed behavioral flexibility seems to go beyond what might be expected if the mind were working with maps, images, or utilities.

If, with Johnson et al., we see conviction narratives as representational units, then it is natural to ask: What representational assumptions do narrative representations embody? Which pieces of knowledge or types of decisions should it not be possible to represent (or represent easily) using narratives?

Taking the narrative story at face value, we might expect that people should not be able to represent narratives which have "plot holes" or contradictions – precisely because they should not have a natural story-like representation. Yet, as with impossible figures, plot holes and contradictions often go undetected. In *The Big Sleep*, the murder of the butler is famously both unexplained, and seemingly inexplicable – Raymond Chandler admitted that he himself had no idea who was responsible (Herman, 1997). Thus, a gaping plot hole evaded both many readers and the author. But the same point arises, of course, for much simpler cases. The "Moses illusion" (Erickson & Mattson, 1981) asks people how many of each animal did Moses take onto the ark; "two" is often the ready answer, although, of course, there is no biblical or other story in which Moses took any animals onto an ark (it was, of course, Noah)

These sorts of cases aren't necessarily problematic for Johnson et al.'s analysis – but I think they would pose a challenge for any model in which narratives form the *building-blocks* of knowledge. Because if narratives are the building-blocks of knowledge, then these fundamental narratives need to be coherent in their own terms.

An alternative viewpoint is that the narratives are not building blocks at all, but are the results of cognitive processing [just as we might see visual imagery and indeed visual perception as a result of symbolic computation, rather than as arising from a distinctively pictorial mode of representation (Pylyshyn, 2002, 2003)]. Rather, we might see the mind is an inveterate story-spinner that continually generates narratives, moment-by-moment, to make sense of the world around us (Chater, 2018). So, for example, consider observing Heider and Simmel's (1944) celebrated animation of an "interaction" in which an aggressive large triangle chases a small triangle and circle. We find ourselves projecting a story-like interpretation, in which the small triangle is initially defending the circle, but is forced aside; the circle "escapes" and locks the large triangle in a confined room (denoted by a rigid rectangle); the circle and small triangle keep themselves hidden; the furious large triangle ends up by "smashing up" the room. Our tendency to populate the world with narratives involving plans, intentions, friendships, and hostilities is, indeed, ubiquitous. But more evidence may be required to justify Johnson et al.'s proposal that narratives play a key role in mental representation.

**Financial support.** This work was supported by the ESRC Network for Integrated Behavioural Science [grant number ES/K002201/1].

**Competing interest.** None.

## References

Bloom, P. (2005). *Descartes' baby: How the science of child development explains what makes us human*. Random House.
Chater, N. (2018). *The mind is flat*. Penguin.
Dennett, D. C. (1978). A cure for the common code. In D. C. Dennett (Ed.), *Brainstorms: Philosophical essays on mind and psychology* (pp. 90–108). Bradford Books.
Erickson, T. A., & Mattson, M. E. (1981). From words to meaning: A semantic illusion. *Journal of Verbal Learning and Verbal Behavior*, *20*, 540–552.
Fodor, J. A. (1975). *The language of thought*. Harvard University Press.
Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *American Journal of Psychology*, *57*(2), 243–259.
Herman, J. (1997). Mysteries of "The Big Sleep". *The Los Angeles Times*, 4 December. https://www.latimes.com/archives/la-xpm-1997-dec-04-ca-60391-story.html
Kosslyn, S. M. (1980). *Image and mind*. Harvard University Press.
Muryy, A., & Glennerster, A. (2021). Route selection in non-Euclidean virtual environments. *PLoS ONE*, *16*(4), e0247818.
Penrose, L. S., & Penrose, R. (1958). Impossible objects: A special type of visual illusion. *British Journal of Psychology*, *49*, 31–33.
Pylyshyn, Z. W. (1981). The imagery debate: Analogue media versus tacit knowledge. *Psychological Review*, *88*(1), 16–45.
Pylyshyn, Z. W. (2002). Mental imagery: In search of a theory. *Behavioral and Brain Sciences*, *25*(2), 157–182.
Pylyshyn, Z. W. (2003). *Seeing and visualizing: It's not what you think*. MIT Press.
Schank, R. C., & Abelson, R. P. (1977). *Scripts, plans, goals and understanding: An inquiry into human knowledge structures*. Lawrence Erlbaum Associates.
Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, *55*(4), 189–208.
Tsetsos, K., Moran, R., Moreland, J., Chater, N., Usher, M., & Summerfield, C. (2016). Economic irrationality is optimal during noisy decision making. *Proceedings of the National Academy of Sciences*, *113*, 3102–3107.

# Narrative thought and decision-making

Peter Ford Dominey

Robot Cognition Laboratory, INSERM U1093 – CAPS, Université de Bourgogne, Dijon Cedex 21078, France.
peter.dominey@inserm.fr
https://www.researchgate.net/profile/Peter-Dominey

doi:10.1017/S0140525X22002692, e90

**Abstract**

A significant body of literature has identified how narrative provides a basis for perceiving and understanding human experience. In the target article, the authors arrive at the need for a form of narrative-based reasoning due to constraints that render probabilistic-based reasoning ineffective. This commentary attempts to bridge this gap and identify links between the proposed and existing theories.

Johnson et al. expose a clear example of decision-making in the face of radical uncertainty that can rely on narrative practice. Their approach is all the more interesting as it has come into being entirely independent of the highly analogous work in the narrative tradition of Bruner and Vygotsky, Hutto and Gallagher, Nelson and others (Bruner, 1991; Gallagher & Hutto, 2008; Nelson, 2009; Ricoeur, 1984; Vygotsky, 1965), and thus provides an unbiased supporting case for these theories. In the following I will first briefly situate how this related work on narrative practice and narrative thinking is potentially pertinent to the authors' work, and then I will similarly review work on narrative and situation models in the context of the authors' proposal for narrative simulation. The hope is that Conviction Narrative Theory (CNT) can be strengthened by building on these foundations, and that reciprocally it can provide a case study for these narrative theories.

Under the heading of "Two modes of thought," Jerome Bruner argued that in addition to rational thought based on logical reasoning, there is a perhaps more important and pervasive mode of human thought based on narrative (Bruner, 1986, 1987). The two modes are related to the distinction between how to know truth versus how to endow experience with meaning. One mode, paradigmatic or logico-scientific, is concerned with establishing empirical truth. The other mode, narrative thinking, deals in human intention and action and their consequences, and making meaning. Bruner attributes a primal importance to narrative in the human drive to attribute meaning to events, and speaks of narrative not only in representing, but actually in constituting reality (Bruner, 1991), and argues that "so our experience of human affairs comes to take the form of the narratives we use in telling about them." This narrative mode of thought, and the need for making sense of human behavior, that is developed by Bruner is quite similar to Johnson et al.'s proposal that narrative provides an alternative to reasoning based on probabilities under radical uncertainty. Thus, Bruner identifies the distinction, and this second mode of thought that is independently proposed in the target article.

This primacy of narrative in human understanding is likewise highlighted by Gergen and Gergen (1988) who insist that stories (narratives) render events understandable and allow expectations for future events. Ricoeur (1984, 1991) further characterizes this human predisposition to organize experience in a narrative format, in terms of agents, goals, actions, etc., the ability to encode experience in a narrative form, and to understand underlying causes in such narrative structures. This infrastructure for understanding human experience in the context of narrative should be of use in CNT.

Hutto and Gallagher exploit this notion of narrative thought for understanding human behavior in their Narrative Practice Hypothesis. This narrative practice theory holds that in human interaction, people regularly generate folk psychological narratives that explain why a person acted on a particular occasion, and that through exposure to these narratives children acquire the skills to understand and themselves produce such narratives (Gallagher, 2013; Gallagher & Hutto, 2008; Hutto, 2007, 2009). Thus, theories of narrative practice hold that through normal exposure to narratives about human social interaction the child will come to learn how to interpret, react to, and respond to social contexts as provided by a theory of mind or folk psychology where psychological decision-making is made based on narrative patterns (Gallagher & Hutto, 2008; Hutto, 2007; Nelson, 2009). Again, this framework should be of interest as it supports and provides a foundation for the hypotheses in CNT.

In a related way, the authors develop ideas about how reasoning about the future can be achieved by "projecting a narrative forward" in section 7.1. While the intention is good, I believe that a more robust way to consider this would be based on the notion of situation model (Zwaan, 2016; Zwaan & Radvansky, 1998). A situation model is a mental representation of the state of affairs denoted by a narrative, rather than only a mental representation of the narrative itself. The situation model is thus a richer higher dimensional representation, and contains information that is not explicit in the narrative itself. In this perspective, the narrative is only a partial representation of the much richer underlying situation model that it describes. For forward prediction, it is thus the situation model itself that provides the basis for simulation (Zwaan, 2016) and projection into the future that can be of service in decision-making.

We recently developed a model and situated robotic system that combines sensorimotor and perceptual inputs and narrative inputs to construct a situation model. Through its own experience and human narration, the system uses narrative practice to build up narrative constructions or patterns that are associated with representations of the causal and temporal structure of physical and mental states in the situation model (Mealier et al., 2017; Pointeau, Mirliaz, Mealier, & Dominey, 2021). In a proof of concept, we demonstrated how the system could use existing narrative constructions to understand (and thus potentially make decisions as required in CNT) about newly experienced situations, inspired by the narrative practice hypothesis of Hutto and Gallagher, and the notion of narrative structure of reality and perception of Bruner, Ricoeur, Gergen and Gergen, and Nelson. Again, such an infrastructure might be usefully employed in providing a basis for certain capabilities identified in Johnson et al.'s CNT.

The multi-disciplinary work described by Johnson et al. provides a rich testing ground for these theories of narrative practice. Indeed, it is quite remarkable that independent of this literature, the authors have arrived at the need for, and initial specification of, a related narrative practice system. In conclusion, Johnson et al. have arrived at the discovery of the value of narrative thinking as a way to overcome constraints imposed by radical uncertainty. It will be of interest to see if and how they reconcile their theory into the existing landscape of narrative thinking, narrative practice, and situation models.

**Financial support.** This research was supported by the ANR France Relance; and Région Bourgogne-Franche-Comté ANER Robotself.

**Competing interest.** None.

## References

Bruner, J. (1987). Life as narrative. *Social Research: An International Quarterly*, *54*, 691–710.
Bruner, J. (1991). The narrative construction of reality. *Critical Inquiry*, *18*(1), 1–21.
Bruner, J. S. (1986). *Actual minds, possible worlds*. Harvard University Press.
Gallagher, S. (2013). A pattern theory of self. *Frontiers in Human Neuroscience*, *7*, 443.
Gallagher, S, & Hutto, D. (2008). Understanding others through primary interaction and narrative practice. *The Shared Mind: Perspectives on Intersubjectivity*, *12*, 17–38.
Gergen, K. J., & Gergen, M. M. (1988). Narrative and the self as relationship. *Advances in Experimental Social Psychology*, *21*, 17–56.
Hutto, D. D. (Ed.). (2009). *Narrative and folk psychology*. Imprint Academic.
Hutto, D. D. (2007). The narrative practice hypothesis: Origins and applications of folk psychology. *Royal Institute of Philosophy Supplement*, *60*, 43–68.
Mealier, A. L., Pointeau, G., Mirliaz, S., Ogawa, K., Finlayson, M., & Dominey, P. F. (2017). Narrative constructions for the organization of self experience: Proof of concept via embodied robotics. *Frontiers in Psychology*, *8*, 1331.
Nelson, K. (2009). Narrative practices and folk psychology: A perspective from developmental psychology. *Journal of Consciousness Studies*, *16*, 69–93.
Pointeau, G., Mirliaz, S., Mealier, A. L., & Dominey, P. F. (2021). Learning to use narrative function words for the organization and communication of experience. *Frontiers in Psychology*, *12*, 591703.
Ricoeur, P. (1984). *Time and narrative* (Vol. *1*). University of Chicago Press. 274 pp.
Ricoeur, P. (1991). Life in quest of narrative. In D. Wood (Ed.), *On Paul Ricoeur: Narrative and interpretation* (pp. 20–33). Routledge.
Vygotsky, L. (1965). *Thought and language*. MIT Press.
Zwaan, R. A. (2016). Situation models, mental simulations, and abstract concepts in discourse comprehension. *Psychonomic Bulletin & Review*, *23*, 1028–1034.
Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*, 162–185.

# Simulation does not just inform choice, it changes choice

Karalyn F. Enz and Diana I. Tamir

Department of Psychology, Princeton University, Princeton, NJ 08540, USA.
kenz@princeton.edu
dtamir@princeton.edu
https://scholar.princeton.edu/kenz
https://psnlab.princeton.edu/

doi:10.1017/S0140525X22002680, e91

**Abstract**

Simulation – imagining future events – plays a role in decision-making. In Conviction Narrative Theory, people's emotional responses to their simulations inform their choices. Yet imagining one possible future also increases its plausibility and accessibility relative to other futures. We propose that the *act* of simulation, in addition to affective evaluation, drives people to choose in accordance with their simulations.

According to Conviction Narrative Theory (CNT), simulating imagined futures allows people to assess how they feel about those futures. These feelings, in turn, inform people's decisions about which future to ultimately pursue. In this account, people simulate by extending the most plausible narrative account of the present into the imagined future. They then use their emotional response – the output of the simulation process – to decide whether to approach or avoid that future (target article). The key here is that the affective evaluation of the simulation drives people's actions: people choose to act in line with imagined futures they feel good about. While we agree that this is one way simulation can affect decision-making, we propose that simulation does more than just give people information about how they feel about their possible choices. We extend the CNT account by considering how the act of simulation itself can *change* choice.

A growing body of experimental literature finds that simulating a specific event amplifies the perceived likelihood of that event. For example, after a person simulates contracting a disease, they think it is more likely they will get that disease in the future (Sherman, Cialdini, Schwartzman, & Reynolds, 1985); the easier the symptoms are to imagine, the more people perceive the disease as likely. The effects of simulation also accumulate with repetition. For example, as people imagine a social interaction over and over, the event seems more and more plausible (Schacter, Benoit, De Brigard, & Szpunar, 2015; Szpunar & Schacter, 2013). This amplification of plausibility occurs after simulating both good and bad interactions, suggesting that one's affective readout of a simulation may occur independently of these consequences. Simulation even has the power to change people's perceptions of the past, convincing them that they experienced imagined past events (Garry, Manning, Loftus, & Sherman, 1996). Together, these findings suggest that simulation increases the plausibility of simulated events, even when it doesn't make the event feel more affectively palatable. It is as if simulating an event paves the way for people to think about the event as real, whether or not they want it to happen.

Simulating one possible future can also block the path to alternative futures. Research on memory has shown that recalling one piece of information can make people forget related information – a phenomenon known as retrieval-induced forgetting (Anderson, Bjork, & Bjork, 1994). For example, if a person studies a list of fruits and then selectively practices apples and grapes, they become less likely to remember oranges or pears later. This suppression can occur during decision-making as well: if people generate detailed evidence supporting one option, it later becomes harder to come up with support for alternative options (Iglesias-Parro & Gómez-Ariza, 2006; Ting & Wallsten, 2011; Weber & Johnson, 2011). Like remembering, mental imagery can induce forgetting (Saunders, Fernandes, & Kosnes, 2009). Simulation may inadvertently narrow people's options by inhibiting the accessibility of non-simulated futures.

Simulation is clearly more than just a passive way to read out affective responses to choice options. By making the imagined future more concrete and blocking the path to other options, simulation paves the way for a person to walk down the simulated path. Indeed, simulation has the power to shift actual behavior such that people become more likely to enact the simulated future: After people simulate helping another person, they become more likely to *actually* engage in prosocial behavior (Gaesser, Shimura, & Cikara, 2020). After people simulate making one choice among several, they become more likely to *actually* pursue that choice (Enz & Tamir, in prep). For example, in controlled lab scenarios, when presented with two snacks to eat, people are more likely to choose the snack that they were randomly assigned to simulate; when presented with two videos to watch, people are more likely to choose the video that they were randomly assigned to simulate. This shift in choice behavior occurs even when people initially like both options equally. This finding translates to consequential decisions outside of the lab, as well. After people simulate an option for an upcoming decision from their own life, they become more likely to choose the option they simulated and less likely to choose any options they did not simulate.

The choice-promoting influence of simulation feeds a positive feedback loop, since people are more likely to simulate options they are initially more likely to choose (Enz & Tamir, in prep). Several lines of work have shown that people begin a decision-making process by considering the most plausible options. For example, people often start with options that reflect the status quo (Weber & Johnson, 2011); options that are valued, moral, or practical (Morris, Phillips, Huang, & Cushman, 2021; Phillips, Morris, & Cushman, 2019); or options that have worked for them in the past (Klein, 1993, 2005). Simulating these already promising options makes them seem even more attractive, further increasing their likelihood of being simulated, and ultimately chosen. This simulation-induced cognitive feedback loop could help to explain how people become more convinced of their narratives and corresponding imagined futures over time.

Rather than merely providing a route to affective evaluation, we propose that the act of simulation has the power to change one's choices. Importantly, simulation can change choices *independent* of affect. The cognitive effects of simulation increase the likelihood that a person will choose the future they simulate not only because they feel good about that imagined future, but because it becomes the easiest future to imagine and pursue.

**Financial support.** This work was supported by NSF grant BCS-1844559 to D. I. T.

**Competing interest.** None.

## References

Anderson, M. C., Bjork, R. A., & Bjork, E. L. (1994). Remembering can cause forgetting: Retrieval dynamics in long-term memory. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *20*(5), 1063–1087. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/7931095

Enz, K. F., & Tamir, D. I. (in prep). Imagine before you decide: Episodic simulation increases the likelihood of choosing the simulated option.

Gaesser, B., Shimura, Y., & Cikara, M. (2020). Episodic simulation reduces intergroup bias in prosocial intentions and behavior. *Journal of Personality and Social Psychology*, *118*(4), 683–705. doi: 10.1037/pspi0000194.

Garry, M., Manning, C. G., Loftus, E. F., & Sherman, S. J. (1996). Imagination inflation: Imagining a childhood event inflates confidence that it occurred. *Psychonomic Bulletin & Review*, *3*(2), 208–214. doi: 10.3758/BF03212420

Iglesias-Parro, S., & Gómez-Ariza, C. J. (2006). Biasing decision making by means of retrieval practice. *The European Journal of Cognitive Psychology*, *18*(6), 899–908. doi: 10.1080/09541440500501189

Klein, G. (1993). A recognition-primed decision (RPD) model of rapid decision making. In G. A. Klein, J. Orasanu, R. Calderwood, & C. E. Zsambok (Eds.), *Decision making in action: Models and methods* (pp. 138–147). Ablex Publishing.

Klein, G. (2005). Recognition-primed decisions. In N. Moray (Ed.), *Ergonomics: Major writings* (Vol. *3*, pp. 271–307). Taylor & Francis.

Morris, A., Phillips, J., Huang, K., & Cushman, F. (2021). Generating options and choosing between them depend on distinct forms of value representation. *Psychological Science*, *32*(11), 1731–1746. doi: 10.1177/09567976211005702

Phillips, J., Morris, A., & Cushman, F. (2019). How we know what not to think. *Trends in Cognitive Sciences*, *23*(12), 1026–1040. doi: 10.1016/j.tics.2019.09.007

Saunders, J., Fernandes, M., & Kosnes, L. (2009). Retrieval-induced forgetting and mental imagery. *Memory & Cognition*, *37*(6), 819–828. doi: 10.3758/MC.37.6.819

Schacter, D. L., Benoit, R. G., De Brigard, F., & Szpunar, K. K. (2015). Episodic future thinking and episodic counterfactual thinking: Intersections between memory and decisions. *Neurobiology of Learning and Memory*, *117*, 14–21. doi: 10.1016/j.nlm.2013.12.008

Sherman, S. J., Cialdini, R. B., Schwartzman, D. F., & Reynolds, K. D. (1985). Imagining can heighten or lower the perceived likelihood of contracting a disease: The mediating effect of ease of imagery. *Personality & Social Psychology Bulletin*, *11*(1), 118–127. doi: 10.1177/0146167285111011

Szpunar, K. K., & Schacter, D. L. (2013). Get real: Effects of repeated simulation and emotion on the perceived plausibility of future experiences. *Journal of Experimental Psychology. General*, *142*(2), 323–327. doi: 10.1037/a0028877

Ting, H., & Wallsten, T. S. (2011). A query theory account of the effect of memory retrieval on the sunk cost bias. *Psychonomic Bulletin & Review*, *18*(4), 767–773. doi: 10.3758/s13423-011-0099-4

Weber, E. U., & Johnson, E. J. (2011). Query theory: Knowing what we want by arguing with ourselves. *The Behavioral and Brain Sciences*, *34*(2), 91–92. doi: 10.1017/s0140525×10002797

# Psychological frameworks augment even classical decision theories

Matthew Charles Ford and John Anderson Kay

St John's College, University of Oxford, Oxford, OX1 3JP, UK.
matthewcford@icloud.com
johnkay@johnkay.com
www.johnkay.com

doi:10.1017/S0140525X22002709, e92

**Abstract**

Johnson, Bilovich, and Tuckett set out a helpful framework for thinking about how humans make decisions under radical uncertainty and contrast this with classical decision theory. We show that classical theories assume so little about psychology that they are not necessarily in conflict with this approach, broadening its appeal.

Johnson, Bilovich, and Tuckett set out a framework for thinking about how humans make decisions under radical uncertainty. We confront radical uncertainty when there is no objective basis for attaching probabilities to different outcomes of decisions or where these outcomes are themselves subject to vagueness or ambiguity. They explicitly contrast their framework to "classical decision theory, [where] the currency of thought is probability and the driver of action is expected utility maximization" (target article, Fig. 1 legend). This description of classical decision theory is common but inaccurate.

Classical decision theory presents utility maximisation as a representation that can be made given the relevant axioms are met, not as the criteria by which choices are actually made. It is not so much a theory *of* decision-making as a theory *about* decision-making. In particular, these theories take it as axiomatic that the individual already knows what they want to do in every situation. The other axioms do not influence their choices but only constrain them.

Expected Utility Theory's (von Neumann & Morgenstern, 1953) axiom (3:A:a) is "the statement of the completeness of the system of individual preferences" over lotteries, which are "combinations of events with stated [i.e., objective] probabilities." Savage (1954) developed a similar theory for subjective beliefs about probabilities.

Savage's axiom P1 states that there is "a simple ordering among acts." Pfanzagl's (1967) order axiom similarly assumes a complete order over wagers. Anscombe and Aumann (1963) allow for both objective and subjective probabilities and state "we share with him [Savage] explicitly P1." Schmeidler (1989) allows subjective non-additive probabilities; his first axiom is a weak order over all acts. In all these utility theories maximising some function of utility is not "the driver of action"; decision-makers are assumed to already know what they want to do in any situation. We have criticised this assumption elsewhere (Kay & King, 2020) and Aumann (1962) shows formally how these theories become silent when it is dropped.

In contrast, Prospect Theory (Kahneman & Tversky, 1979) is represented as a theory of how decisions are made; prospects are valued by combining a decision weight (a function of objective or subjective probability) and the subjective value attached to a change from a reference level for each potential outcome. Although decision-makers are assumed to be able to value any prospect through this mechanism, Prospect Theory purports to explain how these values are reached, rather than merely representing them. (One could also claim that Prospect Theory is just a representation of an even more complex process.)

Many other theories of decision-making claim to deduce, or perhaps recommend, courses of action by assuming a particular goal for the decision-maker. Markowitz (1952) asserts that an optimal portfolio is constructed by finding the optimal point (for a subjective but consistently employed risk tolerance) on an efficient frontier of return-variance tradeoffs; various growth-optimal approaches assume the growth rate of wealth will be maximised. Elsewhere we have shown that in real situations growth optimality requires further psychological assumptions to actually predict behaviour (Ford & Kay, 2022). Still other theories, such as minimax, assume an even simpler goal, for example, choose such that the worst-possible outcome is as good as it can be.

Still other theories are explicitly behavioural and thus describe only particular observed characteristics of real decision-making. For example, the anchoring effect – exposure to a particular, seemingly irrelevant number influences a subsequent decision – seems to demand a psychological explanation. Many such "biases" (or apparent violations of the axioms of subjective utility theory, as they are often defined in contrast to the predictions of a utility theory) have been posited.

With this taxonomy (Table 1), we see that the central economic theory of decision-making under uncertainty draws on

**Table 1** (Ford and Kay). A comparison of different theories of decision-making under uncertainty

| Theory | Sources | Brief Description |
|---|---|---|
| Expected Utility Theory | von Neumann and Morgenstern (1953) | Objective probabilities; assumes completeness |
| Subjective Expected Utility Theory | Savage (1954); Anscombe and Aumann (1963); Pfanzagl (1967) | Subjective probabilities; assumes completeness |
| Non-Additive Subjective Expected Utility Theory | Schmeidler (1989) | Subjective non-additive probabilities; assumes completeness |
| Prospect Theory | Kahneman and Tversky (1979) | Decision weights (a function of probabilities); assumes preferences over certain outcomes but constructs them over prospects |
| Variance Aversion | Markowitz (1952) | Objective probabilities; assumes a specific goal given risk appetite |
| Growth-Optimality | Summarised in Ford and Kay (2022) | Objective probabilities; assumes a specific goal |
| Focal Values | For example Wald (1945) | Possible outcomes; commonly assumes a specific goal |
| Behavioural/Heuristic | Tversky and Kahneman (1974) | Specific adjustments to decisions relative to other theories; not a complete theory |

no theory of psychology (beyond the naïve "people already know exactly what they want"). Given that assumption, the theory merely shows that, with certain restrictions, decision-making can be expressed in a mathematically tractable form. Other theories, often more influenced by psychology, do introduce assumptions about how decisions are actually made, albeit often retaining the assumptions that outcomes can be exhaustively listed and assigned probability-type measures.

Thus the authors do both less and more than it appears. Less, in that their account of affective evaluation of outcomes is not so far removed from the theories they critique; more, in that by providing a psychological framework for how such evaluation occurs, they are providing an explanation for a process that some of those theories assume but are silent upon, and others cannot fully account for.

**Financial support.** This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

**Competing interest.** None.

## References

Anscombe, F. J., & Aumann, R. J. (1963). A definition of subjective probability. *The Annals of Mathematical Statistics*, *34*(1), 199–205.

Aumann, R. J. (1962). Utility theory without the completeness axiom. *Econometrica*, *30* (3), 445–462. doi: 10.2307/1909888

Ford, M. C., & Kay, J. A. (2022). Psychology is fundamental: The limitations of growth-optimal approaches to decision making under uncertainty. Retrieved from https://dx.doi.org/10.2139/ssrn.4140625

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–292. doi: 10.2307/1914185

Kay, J. A., & King, M. A. (2020). *Radical uncertainty: Decision-making for an unknowable future*. The Bridge Street Press.

Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, *7*(1), 77–91. doi: 10.2307/2975974

Pfanzagl, J. (1967). Subjective probability derived from the Morgenstern-von Neumann utility concept. In M. Shubik (Ed.), *Essays in mathematical economics in honor of Oskar Morgenstern* (pp. 237–251). Princeton University Press.

Savage, L. J. (1954). *The foundations of statistics*. John Wiley and Sons.

Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica*, *57*(3), 571–587. doi: 10.2307/1911053

Tversky, A., & Kahneman, D. (1974) Judgment under uncertainty: Heuristics and biases. *Science*, *185*(4157), 1124–1131.

von Neumann, J., & Morgenstern, O. (1953). *The theory of games and economic behavior* (3rd ed.). Princeton University Press.

Wald, A. (1945). Statistical decision functions which minimize the maximum risk. *Annals of Mathematics*, *46*(2), 265–280. doi: 10.2307/1969022

# Really radical?

Karl Friston

The Wellcome Centre for Human Neuroimaging, UCL Queen Square Institute of Neurology, London WC1N 3AR, UK.
k.friston@ucl.ac.uk
https://www.fil.ion.ucl.ac.uk/~karl/

doi:10.1017/S0140525X2200276X, e93

**Abstract**

I enjoyed reading this compelling account of Conviction Narrative Theory (CNT). As a theoretical neurobiologist, I recognised – and applauded – the tenets of CNT. My commentary asks whether its claims could be installed into a Bayesian mechanics of decision-making, in a way that would enable theoreticians to model, reproduce and predict decision-making.

Conviction Narrative Theory (CNT) (target article) is both a "theory of narratives" and a "narrative theory" that precludes mathematical or numerical analysis. This commentary reviews the commitments of CNT through the lens of active inference and self-evidencing (Hohwy, 2016), asking whether CNT could lend itself to a formal (Bayesian) treatment.

Box 1 summarises the fundaments of active inference, as it relates to decision-making under uncertainty. With these fundaments, one can simulate the kind of decision-making addressed by CNT. For example, active inference reproduces decision-making under unknowable circumstances (Friston et al., 2016); it dissolves the exploration–exploitation dilemma and provides a principled account of affordances (Schwartenbeck et al., 2019). It can model the spread of ideas (Albarracin, Demekas, Ramstead, & Heins, 2022) and has been applied to cultural niche construction and social norms (Veissiere, Constant, Ramstead, Friston, & Kirmayer, 2019).

What prevents CNT from using active inference for simulation, scenario modelling or computational phenotyping (Parr, Rees, & Friston, 2018)? One answer is that the requisite generative models are too complex and difficult to specify. However, there may be some commitments of CNT that could be usefully

**Box 1** (Friston). Active inference

Recent trends in theoretical neurobiology, machine learning and artificial intelligence converge on a single imperative that explains both sense-making and decision-making in self-organising systems, from cells (Friston, Levin, Sengupta, & Pezzulo, 2015) to cultures (Veissiere et al., 2019). This imperative is to maximise the evidence (a.k.a. marginal likelihood) for generative (a.k.a., world) models of how observations are caused. This imperative can be expressed as minimising an evidence bound called variational free energy (Winn & Bishop, 2005) that comprises complexity and accuracy (Ramstead et al., 2022):

$$\text{Free energy} = \text{model complexity} - \text{model accuracy}$$

Accuracy corresponds to goodness of fit, while complexity scores the divergence between prior beliefs (before seeing outcomes) and posterior beliefs (afterwards). In short, complexity scores the information gain or cost of changing one's mind (in an information theoretic and thermodynamic sense, respectively). This means Bayesian belief updating is about finding an accurate explanation that is minimally complex (c.f., Occam's principle). In an enactive setting – apt for explaining decision-making – beliefs about "which plan to commit to" are based on the free energy expected under a plausible plan. This implicit planning as inference can be expressed as minimising expected free energy (Friston, Daunizeau, Kilner, & Kiebel, 2010):

$$\text{Expected free energy} = \text{risk (expected complexity)} + \text{ambiguity (expected inaccuracy)}$$

Risk is the divergence between probabilistic predictions about outcomes, given a plan, relative to prior preferences. Ambiguity is the expected inaccuracy. An alternative decomposition is especially interesting from the perspective of CNT:

$$\text{Expected free energy} = \text{expected cost} - \text{expected information gain}$$

The expected information gain underwrites the principles of optimal Bayesian design (Lindley, 1956), while expected cost underwrites Bayesian decision theory (Berger, 2011). However, there is a twist that distinguishes active inference from expected utility theory. In active inference, there is no single, privileged outcome that furnishes a utility or cost function. Rather, utilities are replaced by preferences, quantified by the (log) likelihood of encountering *every* aspect of an observable outcome. In short, active inference appeals to two kinds of Bayes optimality and subsumes information and preference-seeking behaviour under a single objective.

dismantled, enabling its claims to be substantiated with simulations and implicit proof of principle.

In active inference, narratives feature as prior beliefs. Indeed, the plans – that underwrite policy selection – are often described as narratives (Friston, Rosch, Parr, Price, & Bowman, 2017b). So, could one cast CNT in terms of narrative (i.e., policy) selection and "planning as inference" (Attias, 2003; Botvinick & Toussaint, 2012; Matsumoto & Tani, 2020)? In what follows, five arguments against formalising CNT in this fashion are considered and countered.

(1) Radical uncertainty does not admit any Bayesian mechanics because the requisite probabilities do not have a well-defined outcome space.

Radical uncertainty rests upon an unknowable *outcome* (e.g., John Kay's wheel example). However, outcomes are *known* quantities that are observed. Technically, radical uncertainty refers to unknowable (i.e., hidden) *causes* of outcomes. However, finding the right causal explanation just is the problem of Bayesian inference. So what is radical about radical uncertainty? The answer might lie in the hierarchical nature of belief updating and implicit generative models. Given the parameters of a generative model, I can be uncertain about hidden *states* generating my observations. However, I can also be uncertain about the *parameters*, given a model. Finally, I can have uncertainty about my *model*. Radical uncertainty seems to concern the model structure.

Resolving the three kinds of uncertainty above corresponds to *inference*, *learning* and *model selection*, respectively. All entail maximising marginal likelihood or minimising free energy, with respect to posteriors over states, parameters and models, respectively. Model selection is known as structure learning in Radical Constructivism (Salakhutdinov, Tenenbaum, & Torralba, 2013; Tenenbaum, Kemp, Griffiths, & Goodman, 2011; Tervo, Tenenbaum, & Gershman, 2016). Structure learning is a partly solved problem, through *Bayesian model reduction* (Friston, Parr, & Zeidman, 2018), where redundant components are removed from an overly expressive model to maximise model evidence (e.g., Smith, Schwartenbeck, Parr, & Friston, 2020). This reductive approach complements non-parametric Bayes, which formalises the inclusion of new narratives (Gershman & Blei, 2012). In light of Bayesian model selection, one could argue that radical uncertainty admits a Bayesian mechanics.

(2) The utilities of different kinds of outcomes cannot be compared in a meaningful way.

This is only a problem if one subscribes to Bayesian decision theory as a complete account. Active inference vitiates this objection because to be Bayes optimal is to resolve uncertainty in the context of securing preferred outcomes (i.e., minimise expected free energy; see Box 1). Crucially, expected utility and information gain (Howard, 1966; Kamar & Horvitz, 2013; Moulin & Souchay, 2015) share the same currency; namely, natural units (when using natural logarithms of prior preferences). This lends a quantitative and comparable meaning to the value of information and preferences.

(3) Certain outcomes are so fuzzy they are impossible to predict and therefore one has to use heuristics.

Knowing something is unpredictable is itself an informative prior that can be installed into hierarchal generative models: c.f., Jaynes' maximum entropy principle (Jaynes, 1957; Kass & Raftery, 1995; Sakthivadivel, 2022). So, how do "fast and frugal" heuristics fit into active inference? Heuristics are generally considered as priors that comply with complexity minimising imperatives (Box 1), for example, habitisation (FitzGerald, Dolan, & Friston, 2014) or the minimisation of perceptual prediction errors (Mansell, 2011). In short, heuristics are exactly what active inference – under

https://doi.org/10.1017/S0140525X22001157 Published online by Cambridge University Press

hierarchal generative models – is there to explain. On this reading of active inference, self-evidencing just is *satisfising* (Gerd Gigerenzer, personal communication).

(4) But people don't behave as if they were rational, or even with bounded rationality.

Many careful studies in cognitive neuroscience are concerned with how people deviate from Bayes optimality. However, this overlooks the complete class theorem (Brown, 1981; Wald, 1947). The complete class theorem says that for any pair of choice behaviours and cost functions, there are some priors that render the decisions Bayes-optimal. This has the fundamental implication that Bayesian mechanics cannot *prescribe* optimal (i.e., rational) decision-making. It can only *describe* rationality in terms of the priors a subject brings to the table. This insight underwrites the emerging field of computational psychiatry, where the game is to estimate the prior beliefs of patients that best explain their decision-making (Schwartenbeck & Friston, 2016; Smith, Khalsa, & Paulus, 2021).

(5) But the dimensionality and numerics of belief updating in realistic generative models are beyond the capacity of any computer, human or otherwise.

This argument rests on the use of sampling procedures to approximate posterior distributions, for example, likelihood free methods or approximate Bayesian computation (Chatzilena, van Leeuwen, Ratmann, Baguelin, & Demiris, 2019; Cornish & Littenberg, 2007; Girolami & Calderhead, 2011; Ma, Chen, & Fox, 2015; Silver & Veness, 2010). However, active inference rests on variational schemes found in physics, high-end machine learning (Marino, 2021) and (probably) the brain (Friston, Parr, & de Vries, 2017a). Variational Bayes eschews sampling by committing to a functional form for posterior beliefs; thereby converting an impossible marginalisation problem into an optimisation problem; namely, minimising variational free energy (Feynman, 1972). In summary, some people may think generative models with realistic narratives cannot be inverted; however, they (i.e., these people) are are existence proofs that such models can be inverted.

**Financial support.** KF is supported by funding for the Wellcome Centre for Human Neuroimaging (Ref: 205103/Z/16/Z) and a Canada-UK Artificial Intelligence Initiative (Ref: ES/T01279X/1).

**Competing interest.** None.

## References

Albarracin, M., Demekas, D., Ramstead, M. J. D., & Heins, C. (2022). Epistemic communities under active inference. *Entropy*, *24*(4).
Attias, H. (2003). Planning by probabilistic inference, Proc. of the 9th Int. Workshop on Artificial Intelligence and Statistics.
Berger, J. O. (2011). *Statistical decision theory and Bayesian analysis*. Springer.
Botvinick, M., & Toussaint, M. (2012). Planning as inference. *Trends in Cognitive Sciences*, *16*, 485–488.
Brown, L. D. (1981). A complete class theorem for statistical problems with finite-sample spaces. *Annals of Statistics*, *9*, 1289–1300.
Chatzilena, A., van Leeuwen, E., Ratmann, O., Baguelin, M., & Demiris, N. (2019). Contemporary statistical inference for infectious disease models using Stan. *Epidemics*, *29*, 100367.
Cornish, N. J., & Littenberg, T. B. (2007). Tests of Bayesian model selection techniques for gravitational wave astronomy. *Physical Review D*, *76*(8).
Feynman, R. P. (1972). *Statistical mechanics*. Benjamin.
FitzGerald, T., Dolan, R., & Friston, K. (2014). Model averaging, optimal inference, and habit formation. *Frontiers in Human Neuroscience*, *8*, 457. doi: 10.3389/fnhum.2014.00457 [KF]
Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., O'Doherty, J., & Pezzulo, G. (2016). Active inference and learning. *Neuroscience and Biobehavioral Reviews*, *68*, 862–879.
Friston, K., Levin, M., Sengupta, B., & Pezzulo, G. (2015). Knowing one's place: A free-energy approach to pattern regulation. *Journal of the Royal Society, Interface*, *12*, 105.
Friston, K., Parr, T., & de Vries, B. (2017a). The graphical brain: Belief propagation and active inference. *Network Neuroscience*, *1*, 381–414.
Friston, K., Parr, T., & Zeidman, P. (2018). Bayesian model reduction. arXiv preprint arXiv:1805.07092.
Friston, K. J., Daunizeau, J., Kilner, J., & Kiebel, S. J. (2010). Action and behavior: A free-energy formulation. *Biological Cybernetics*, *102*, 227–260.
Friston, K. J., Rosch, R., Parr, T., Price, C., & Bowman, H. (2017b). Deep temporal models and active inference. *Neuroscience and Biobehavioral Reviews*, *77*, 388–402.
Gershman, S. J., & Blei, D. M. (2012). A tutorial on Bayesian nonparametric models. *Journal of Mathematical Psychology*, *56*, 1–12.
Girolami, M., & Calderhead, B. (2011). Riemann manifold Langevin and Hamiltonian Monte Carlo methods. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *73*(2), 123–214.
Hohwy, J. (2016). The self-evidencing brain. *Noûs*, *50*, 259–285.
Howard, R. (1966). Information value theory. *IEEE Transactions on Systems, Science and Cybernetics SSC*, *2*, 22–26.
Jaynes, E. T. (1957). Information theory and statistical mechanics. *Physical Review Series II*, *106*, 620–630.
Kamar, E., & Horvitz, E. (2013). Light at the end of the tunnel: a Monte Carlo approach to computing value of information. In Proceedings of the 2013 International Conference on Autonomous Agents and Multi-agent Systems, pp. 571–578.
Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*, 773–795.
Lindley, D. V. (1956). On a measure of the information provided by an experiment. *Annals of Mathematical Statistics*, *27*, 986–1005.
Ma, Y.-A., Chen, T., & Fox, E. B. (2015). A complete recipe for stochastic gradient MCMC, p. arXiv:1506.04696.
Mansell, W. (2011). Control of perception should be operationalized as a fundamental property of the nervous system. *Topics in Cognitive Science*, *3*, 257–261.
Marino, J. (2021). Predictive coding, variational autoencoders, and biological connections. *Neural Computation*, *34*, 1–44.
Matsumoto, T., & Tani, J. (2020). Goal-directed planning for habituated agents by active inference using a variational recurrent neural network. *Entropy*, *22*, 564.
Moulin, C., & Souchay, C. (2015). An active inference and epistemic value view of metacognition. *Cognitive Neuroscience*, *6*, 221–222.
Parr, T., Rees, G., & Friston, K. J. (2018). Computational neuropsychology and Bayesian inference. *Frontiers in Human Neuroscience*, *12*, 61.
Ramstead, M. J. D., Sakthivadivel, D. A. R., Heins, C., Koudahl, M., Millidge, B., Da Costa, L., … Friston, K. J. (2022). On Bayesian mechanics: A physics of and by beliefs, p. arXiv:2205.11543.
Sakthivadivel, D. A. R. (2022). A constraint geometry for inference and integration, p. arXiv:2203.08119.
Salakhutdinov, R., Tenenbaum, J. B., & Torralba, A. (2013). Learning with hierarchical-deep models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*, 1958–1971.
Schwartenbeck, P., & Friston, K. (2016). Computational phenotyping in psychiatry: A worked example. *eNeuro*, *3*, 0049-16.2016.
Schwartenbeck, P., Passecker, J., Hauser, T. U., FitzGerald, T. H., Kronbichler, M., & Friston, K. J. (2019). Computational mechanisms of curiosity and goal-directed exploration. *eLife*, *8*, e41703.
Silver, D., & Veness, J. (2010). Monte-Carlo planning in large POMDPs. Proceedings of the Conference on Neural Information Processing Systems.
Smith, R., Khalsa, S. S., & Paulus, M. P. (2021). An active inference approach to dissecting reasons for nonadherence to antidepressants. *Biological Psychiatry-Cognitive Neuroscience and Neuroimaging*, *6*, 919–934.
Smith, R., Schwartenbeck, P., Parr, T., & Friston, K. J. (2020). An active inference approach to modeling structure learning: Concept learning as an example case. *Frontiers in Computational Neuroscience*, *14*, 41.
Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to grow a mind: Statistics, structure, and abstraction. *Science*, *331*, 1279–1285.
Tervo, D. G., Tenenbaum, J. B., & Gershman, S. J. (2016). Toward the neural implementation of structure learning. *Current Opinion in Neurobiology*, *37*, 99–105.
Veissiere, S. P. L., Constant, A., Ramstead, M. J. D., Friston, K. J., & Kirmayer, L. J. (2019). Thinking through other minds: A variational approach to cognition and culture. *The Behavioral and Brain Sciences*, *43*, e90.

Wald, A. (1947). An essentially complete class of admissible decision functions. *The Annals of Mathematical Statistics*, *18*, 549–555.
Winn, J., & Bishop, C. M. (2005). Variational message passing. *Journal of Machine Learning Research*, *6*, 661–694.

# How do narratives relate to heuristics?

Gerd Gigerenzer[a,b]

[a]Center for Adaptive Behavior & Cognition, Max Planck Institute for Human Development, Berlin 14193, Germany and [b]Harding Center for Risk Literacy, University of Potsdam, Potsdam, Germany.
gigerenzer@mpib-berlin.mpg.de

doi:10.1017/S0140525X22002710, e94

**Abstract**

Narratives and heuristics are both tools for dealing with uncertainty, intractability, and incommensurability, that is, for all real-world situations outside the domain of Bayesian decision theory. But how do narratives and heuristics relate? I suggest two links: Heuristics select narratives to explain events, and "big" narratives select the heuristics that people live by, to execute their values and moral principles.

The tools of Bayesian decision theory – the consistency axioms, maximization of expected utility, and Bayesian probability updating – are tailored for situations of risk, not uncertainty. In Savage's (1972/1954) terms, a situation of risk means a *small world* (S, C) in which all possible future states S, all possible consequences C, and all probabilities are known. A game of roulette is an example. If the state space (S, C) is not fully known, as in most real-world decisions, the situation is one of (radical) uncertainty, also called a *large world*. Examples include finding the right partner, raising one's children, and running the Bank of England. Savage emphasized that applying Bayesian decision theory to situations of uncertainty "is utterly ridiculous".

Dealing successfully with uncertainty, intractability, and incommensurability – all situations beyond the reach of Bayesian decision theory – requires different tools. Narratives are one tool (target article; Tuckett, 2011), heuristics are another (Gigerenzer, Hertwig, & Pachur, 2011). A narrative provides a story while a heuristic provides a concrete sequence of actions, such as what information to search, when to stop, and how to decide. People use heuristics to find a partner, raise their children, run large institutions, and make investment decisions. As Johnson et al. note, developing a repertoire of heuristics and learning to select these in an adaptive way is often the best one can do in the real world.

How do heuristics and narratives relate to each other? One possibility might be that heuristics select narratives, a second one that narratives select heuristics.

Johnson et al.'s burglar-versus-cat narrative illustrates the first link: An observation is made (hearing a noise at night) and, in the absence of probabilities, heuristics are used to evaluate and select one of the two narratives. Here, Johnson et al. see the role of heuristics akin to hypotheses testing. To me, the burglar-versus-cat story is not the best example to highlight radical uncertainty, while the references to investors, banks, and central banks clearly are. In the burglar-versus-cat narrative, all uncertainty appears to be limited to the probabilities, akin to a form of a small world known as *ambiguity*, and the set of possible states of the world (burglar versus cat) is known. In contrast, the "big" narratives on which religion, politics, and science are based are rarely selected on a daily basis. Rather, these narratives select what we believe and how we behave until a revolution in thought occurs. In other words, big narratives select heuristics.

Consider what Max Weber (2001/1904) called the *Protestant ethic*. Work hard and accumulate capital. Don't waste one's earnings on pleasure, power, or material comfort. Instead, live frugally and reinvest them to accumulate more capital. It is based on the teachings of various Puritan religions, including Calvinism, Methodism, Pietism, and Baptism. According to Weber, the underpinning is the doctrine of predestination: God has already chosen who will be saved from damnation or not. All one can do is to look for cues to find out whether one is among the chosen. Working hard and not wasting time on worldly pleasures is such a clue. Wasting time playing billiards, by contrast, is a sign of being doomed. Benjamin Franklin, a proponent of the Protestant work ethic, coined the term "time is money." To the present day, many people have internalized the heuristics for accumulating capital and not wasting time, including the associated feelings of guilt, without necessarily being aware of the underlying narrative. Narratives can be unconscious, while the associated heuristics actively guide life and moral judgments.

Weber contrasts this ethic with a traditional Catholic narrative, where individual fates are not predetermined and forgiveness is possible. People living by this narrative can play billiards, eat well, and enjoy their life without feeling guilty. Weber tells the story of employers who increased the hourly wages of their workers to make them work longer for a limited time, such as at harvest time. Yet to their surprise, many workers did not work more, but fewer hours. Their narrative selected a satisficing heuristic with an amount of money as the aspiration level. When they had earned enough, they stopped working, went home, and spent their earnings and free time together with their family. To make them work longer would have meant reducing, not increasing, their hourly wages. The Protestant work ethic, in contrast, selects a heuristic that tries to boost capital: Work hard, accumulate as much capital as possible, and do not spend it on pleasure. In both cases, the narrative is first, and it selects the heuristics to execute the narrative.

Narratives can select heuristics, and heuristics can select narratives. These are probably not the only links. Another possibility is that the heuristics we live by are learned by imitation, and a narrative is constructed around these to create a consistent story after the fact, or make sense of what one does. A final possibility is that narratives and heuristics co-evolve; changing one changes the other. Further thinking is needed.

All in all, I congratulate Johnson et al. for having the courage to write about the importance of narratives in the context of decision-making. That stands in stark contrast to the widespread belief that all decisions in the real world can be reduced to small worlds where Bayesian decision-making can find the optimal answer. Savage, known as the father of modern decision theory, warned about this misconception. It renders pointless all of psychology, from causal stories to trust and emotion. The belief that we live in a small world is itself a powerful narrative that selects the research questions asked and the experimental tasks studied. We should make our theories more relevant for large worlds, and to do so, we need to take uncertainty seriously, along with the narratives and heuristics by which we live.

**Financial support.** This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

**Competing interest.** None.

## References

Gigerenzer, G., Hertwig, R., & Pachur, T. (Eds.) (2011). *Heuristics: The foundations of adaptive behavior*. Oxford University Press.
Savage, L. J. (1972/1954). *The foundations statistics*. Wiley.
Tuckett, D. (2011). *Minding the markets: An emotional finance view of financial markets*. Palgrave Macmillan.
Weber, M. (2001). *The Protestant ethic and the spirit of capitalism* (T. Parsons, Trans.). Routledge. (Original work published 1904).

# The small world's problem is everyone's problem, not a reason to favor CNT over probabilistic decision theory

Daniel Greco

Department of Philosophy, Yale University, New Haven, CT 06511-6629, USA.
daniel.greco@yale.edu
https://sites.google.com/site/dlgreco/

doi:10.1017/S0140525X22002771, e95

**Abstract**

The case for the superiority of Conviction Narrative Theory (CNT) over probabilistic approaches rests on selective employment of a double standard. The authors judge probabilistic approaches inadequate for failing to apply to "grand-world" decision problems, while they praise CNT for its treatment of "small-world" decision problems. When both approaches are held to the same standard, the comparative question is murkier.

Johnson, Bilovich, and Tuckett marshal evidence from a wide range of sources to make the case that decision-making under radical uncertainty involves essential use of *narratives*; structured, higher-order representations that include information about the causal, temporal, valence, and analogical structure of the decision problems they represent. They make a compelling case that Conviction Narrative Theory (CNT) deserves a seat at the table in theorizing about decision-making under radical uncertainty. But they go further and claim that CNT is superior to theoretical frameworks in which probabilistic inference and expected utility maximization take center stage. I'll argue that their case rests on selective employment of a double standard. They argue that probabilistic inference is impossible under conditions of radical uncertainty, while the task of selecting the most plausible narrative is possible. In order for the latter task to be possible, however, various heuristics must be used to first winnow down the space of possible narratives to a small set of compact narratives from which the best can tractably be selected. But if such heuristics are allowed to play a role in CNT, they should also be allowed to play a role in probabilistic approaches. Once they are, the asymmetry between CNT and probabilistic approaches with respect to their ability to handle radical uncertainty disappears.

Savage (1954) famously distinguished between "small-world" and "grand-world" decision problems. Small-world decision problems are the ones we encounter in textbooks. In such problems the spaces of possible outcomes, actions, and evidence are highly constrained, and it is possible to calculate which action maximizes expected utility, as well as how this would change were the agent in question to gain various pieces of evidence. Small-world decision problems are tractable because they represent only a tiny subset of the possible distinctions that can be made. Grand-world problems represent an agent facing a vast array of options, considering all logically coherent hypotheses about the outcomes of their choosing those options, and considering the relevance of all their sensory and background information to all of the previous. Nobody has ever written down a grand-world decision problem, and if, *per impossibile*, someone did, they'd still be unable to solve it. Savage admitted he had little to say about how we identify small-world problems for analysis, and effectively treated the process as a black box.

While Savage distinguished between small- and grand-world decision problems in the context of expected utility theory, essentially the same distinction can be made in other paradigms for thinking about rational choice. The idea that grand-world decision problems are intractable is very closely related to the infamous "frame problem" for artificial intelligence (McCarthy & Hayes, 1981), in which it's not assumed that an agent must select an optimal action by computing expected utilities. In this more general setting, the problem is how an agent identifies a representation of her practical situation – one that contains enough relevant information for solving it to be fruitful, while still being compact enough to be tractably analyzable – as the one to subject to some decision-making rule. This problem arises no less for CNT than for probabilistic approaches to decision-making; the space of possible narratives is vast and unstructured, as is the space of possible items of evidence an agent might compare for consilience with one or another narrative.

While the authors acknowledge the size of the space of possible narratives, they have very little to say about how we manage to identify tractable, small-world problems of narrative comparison; their discussion of how we use explanatory heuristics to evaluate narratives effectively assumes we *already* have a small set of candidate narratives on the table for evaluation. This argumentative strategy amounts to a double standard. Most of the authors' arguments for the inapplicability of probabilistic approaches to decision-making under radical uncertainty – arguments about the impossibility of enumerating all possible outcomes of our actions, all relevant pieces of evidence, and of coming up with both prior and conditional probabilities for all of the previous – presuppose that probabilistic methods must be applied directly to grand-world decision problems. But when they discuss CNT, they treat the process of how we get from a grand-world problem to a small-world one as a black box; while they mention a necessary role for heuristics, they have nothing to say about how they work, or why similar heuristics couldn't work for probabilistic approaches.

If we ask whether probabilistic approaches can meet the same standard to which the authors hold CNT – being feasibly applicable not directly to grand-world problems, but instead to small-world problems in which some set of heuristics has *already* identified a tractable set of relevant hypotheses, pieces of evidence, and evaluatively relevant features – the prospects aren't so obviously bleak. It's true that even in small-world problems computing exact posterior probabilities is NP-hard (Cooper, 1990). But it's also well known that there are tractable methods for approximate Bayesian computation (Lintusaari, Gutmann, Dutta, Kaski, & Corander, 2017). It's also true that we generally lack non-arbitrary methods for assigning prior probabilities in small-world problems. But given the fruitfulness of Bayesian, probabilistic models in other areas of cognitive science such as vision (Yuille & Kersten, 2006), where worries about arbitrariness aren't obviously any less applicable, it's premature to assume such challenges can't be met in the case of decision-making.

None of this is to undermine the authors' positive case for CNT. Nor is it to say that *all* of their comparative arguments for the superiority of CNT over probabilistic approaches rely on the double standard I describe above; the discussion of "digitization" in section 7.2 is immune from the criticism I've offered here. Still, argumentative double standards should be avoided; neither CNT nor probabilistic approaches to decision-making seem particularly well-suited to answering the exceedingly difficult question of how we identify small-world problems for analysis. Perhaps CNT provides a better account of how we handle such small-world problems *once we've identified them*, but if so, that claim isn't supported by evidence for the infeasibility of applying probabilistic methods directly to grand-world problems.

**Acknowledgements.** None.

**Financial support.** None.

**Competing interest.** None.

## References

Cooper, G. F. (1990). The computational complexity of probabilistic inference using Bayesian belief networks. *Artificial Intelligence*, *42*(2–3), 393–405.

Lintusaari, J., Gutmann, M. U., Dutta, R., Kaski, S., & Corander, J. (2017). Fundamentals and recent developments in approximate Bayesian computation. *Systematic Biology*, *66* (1), e66–e82.

McCarthy, J., & Hayes, P. J. (1981). Some philosophical problems from the standpoint of artificial intelligence. In B. L. Webber & N. J. Nilsson (Eds.), *Readings in artificial intelligence* (pp. 431–450). Morgan Kaufmann. https://doi.org/10.1016/B978-0-934613-03-3.50033-7

Savage, L. J. (1954). *The foundations of statistics*. Wiley Publications in Statistics.

Yuille, A., & Kersten, D. (2006). Vision as Bayesian inference: Analysis by synthesis? *Trends in Cognitive Sciences*, *10*(7), 301–308.

# The wisdom in the story: Clarifying assumptions about radical uncertainty and reasonableness in narrative judgment

Igor Grossmann, Ethan A. Meyers and Richard P. Eibach

Department of Psychology, University of Waterloo, Waterloo, ON N2L 3G1, Canada.

igrossma@uwaterloo.ca
emeyers@uwaterloo.ca
reibach@uwaterloo.ca
https://igorgrossmann.com/
https://uwaterloo.ca/psychology/people-profiles/richard-eibach

doi:10.1017/S0140525X22002746, e96

**Abstract**

Human lives are radically uncertain. Making sense of such uncertainties is the hallmark of wisdom. Sense-making requires narratives, putting them in the center stage of human everyday decision-making. Yet what if radical uncertainty is a narrative itself? Moreover, do laypeople always consider such narratives irrational? Here we pose these questions to enrich a theory of choice under uncertainty.

Everyday human decision-making is neither purely analytic and probability-driven nor is it about whimsical constraint satisfaction. There is nothing radical about this claim – most of the world's philosophical traditions put context-sensitive *meaning-making* instead of logical deduction of correct probabilities at the heart of discourse about what makes up human wisdom (Grossmann, 2017). After all, the ecology of everyday life is rarely well-defined: Preferences of other decision-makers are rarely transparent, and situations typically evolve. Consequently, logically derived probabilities may be of little use for decision-making in the ill-defined contexts of everyday life. Instead, at least since Aristotle scholars have emphasized meta-cognitive faculties that help one to navigate life's uncertainties, get a better sense of the issue at hand (Grossmann, 2017; Grossmann et al., 2020b), and to justify their choices (Grossmann, Eibach, Koyama, & Sahi, 2020a).

In the Conviction Narrative Theory (CNT), Johnson et al. echo these philosophical themes, leveraging insights from cognitive, affective, and behavioral sciences to shed light on how narratives guide human choices under "radical" uncertainty. They postulate a powerful and insightful account of narrative reasoning and the conditions under which people are likely to use it. We applaud the authors' ambitious attempt to integrate disparate bodies of literature in decision sciences and philosophy. At the same time, for CNT to become a common vocabulary in decision sciences, it may need greater clarity about two of its foundational components: Radical uncertainty and reasonableness (versus rationality) as a criterion of good judgment.

CNT postulates that people resort to narrative reasoning in situations where they face radical uncertainty, in which the probabilities of outcomes (or evaluative criteria) are unknowable or ambiguous. Below, we outline how radical uncertainty and fuzzy evaluation may themselves sometimes be the product of narrative reasoning. The person's causal models of the world can determine whether they perceive a given decision problem as a situation of manageable uncertainty, where reasonable estimates of probabilities and values can be assigned to decision outcomes, or a situation where the probabilities of outcomes are too radically uncertain or fuzzy to evaluate using probabilistic reasoning.

Causal models of the world can determine how much uncertainty is perceived in any given decision problem because such models determine what range of factors need to be accounted for. Some folk theories may lead people to situate a particular problem within a relatively tractable deterministic, linear casual system where uncertainty is calculable. For example, Enlightenment thought conceived of a world whose properties were rationally ordered and, at least in principle, knowable. By contrast, other folk theories may situate the same problem within a dynamic, non-linear system in which radical uncertainty applies. For example, postmodern narratives conceive of a world in which most or all knowledge claims are fundamentally suspect and subject to shifting power dynamics within society. Thus, radical uncertainty is the exception within the Enlightenment worldview but the norm within the postmodern worldview.

Notably, there are individual and cultural differences in people's acknowledgement of the uncertainty and changeability of events and awareness of limitations in their knowledge and understanding, which are the core epistemic premises of wise reasoning (e.g., Grossmann et al., 2012; Porter et al., 2022; Santos, Huynh, & Grossmann, 2017). Such differences in lay epistemic assumptions will mean that some decision-makers construe a particular judgment as a situation of radical uncertainty that requires narrative reasoning approaches while others construe it as a situation of manageable uncertainty in which probabilistic reasoning is applicable. Moral narratives may also frame probabilistic reasoning as taboo in certain contexts, such as making decisions involving sacred values (Baron & Spranca, 1997; Tetlock, Kristel, Elson, Green, & Lerner, 2000). Thus, even probabilistic reasoning is ultimately contingent on prior narratives – namely, narratives that frame the judgment as one in which there was relevant, reliable, sufficient information to assess and evaluate the probabilities of outcomes (Douglas, 2001).

Moreover, individuals may be tactically motived to construe a situation as if involving radical uncertainty. In a recent example, even after there was compelling clinical evidence of the efficacy of COVID mRNA vaccines for reducing COVID illness, some vaccine skeptics framed vaccination as a context of radical uncertainty by emphasizing that until long-term data on vaccination outcomes were available it was impossible to know whether vaccination is in one's interest (Lu, 2022). Some people expressed hesitancy about getting a COVID mRNA vaccine due to a variety of imagined long-term side effects ranging from cancer to infertility (Pertwee, Simas, & Larson, 2022). Radical uncertainty can also be rooted in conspiracy narratives that discount official information about probable outcomes as fabrications designed to manipulate the population for nefarious purposes, such as the conspiracy theory that public health information about vaccination efficacy was an elaborate hoax to get tracking microchips into people's bodies (Pertwee et al., 2022). These examples illustrate how radical uncertainty can be a subjective product of narrative reasoning rather than an objective description of the available information.

Our second point concerns criteria of good judgment within CNT. CNT is presented as a type of bounded rationality where the benchmark for sound judgment is reasonableness rather than rationality. Hereby, the definition of reasonableness appears to rely on folk concepts (Grossmann et al., 2020a), whereas the definition of rationality is a smorgasbord of theoretical positions not derived from laypeople. Johnson et al. argue that most definitions of rationality rely on assigning probabilities, which are absent under radical uncertainty and thus disqualify rationality as a possible benchmark.

This argument assumed that the definitions of theorists and of laypeople are equivalent, which is not necessarily so. Lay people define rationality as reductionist and seeking to serve a single value whereas they define reasonableness as holistic and seeking to balance incommensurable values (Meyers, Eibach, Hanxiao, & Grossmann, 2022). The lay definition of rationality does not require probabilities to be assigned, sidestepping the original rejection of rationality as a benchmark. If the lay standard of reasonableness is used for CNT, then the lay standard of rationality should be used, too. This would parallel non-radical decision scenarios wherein people hold both standards to be descriptions of good judgment, dependent on the goal(s) of the decision-maker.

**Financial support.** The present research was funded by the Social Sciences and Humanities Research Council of Canada (Insight grant 435-2014-0685 to I. G.), by a postgraduate scholarship-doctoral grant PGSD3-547482-2020 from the Natural Sciences and Engineering Research Council of Canada (to E. A. M.), by an Early Researcher award ER16-12-169 from the Ontario Ministry of Research and Innovation (to I. G.), by the John Templeton Foundation (grant 62260 to I. G.), and by the Templeton World Charity Foundation (grant TWCF0355 to I. G.).

**Competing interest.** None.

## References

Baron, J., & Spranca, M. (1997). Protected values. *Organizational Behavior and Human Decision Processes*, *70*(1), 1–16. https://doi.org/10.1006/obhd.1997.2690

Douglas, M. (2001). Dealing with uncertainty. *Ethical Perspectives*, *8*(3), 145–155.

Grossmann, I. (2017). Wisdom in context. *Perspectives on Psychological Science*, *12*(2), 233–257. https://doi.org/10.1177/1745691616672066

Grossmann, I., Eibach, R. P., Koyama, J., & Sahi, Q. (2020a). Folk standards of sound judgment: Rationality versus reasonableness. *Science Advances*, *6*(2), eaaz0289. https://doi.org/10.1126/sciadv.aaz0289

Grossmann, I., Karasawa, M., Izumi, S., Na, J., Varnum, M. E., Kitayama, S., & Nisbett, R. E. (2012). Aging and wisdom: Culture matters. *Psychological Science*, *23*(10), 1059–1066. https://doi.org/10.1177/0956797612446025

Grossmann, I., Weststrate, N. M., Ardelt, M., Brienza, J. P., Dong, M., Ferrari, M., … Vervaeke, J. (2020b). The science of wisdom in a polarized world: Knowns and unknowns. *Psychological Inquiry*, *31*(2), 103–133. https://doi.org/10.1080/1047840X.2020.1750917

Lu, J. G. (2022). Two large-scale global studies on COVID-19 vaccine hesitancy over time: Culture, uncertainty avoidance, and vaccine side-effect concerns. *Journal of Personality and Social Psychology*. Advance online publication . https://doi.org/10.1037/pspa0000320

Meyers, E. A., Eibach, R., Hanxiao, Z., & Grossmann, I. (2022). What is good judgment? Descriptive norms and intuitive use of rationality versus reasonableness. *PsyArXiv*. https://doi.org/10.31234/osf.io/fc8n6

Pertwee, E., Simas, C., & Larson, H. J. (2022). An epidemic of uncertainty: Rumors, conspiracy theories and vaccine hesitancy. *Nature Medicine*, *28*, 456–459. https://doi.org/10.1038/s41591-022-01728-z

Porter, T., Elnakouri, A., Meyers, E. A., Shibayama, T., Jayawickreme, E., & Grossmann, I. (2022). Predictors and consequences of intellectual humility. *Nature Reviews Psychology*, *1*, 524–536. https://doi.org/10.1038/s44159-022-00081-9

Santos, H. C., Huynh, A. C., & Grossmann, I. (2017). Wisdom in a complex world: A situated account of wise reasoning and its development. *Social and Personality Psychology Compass*, *11*(1), e12341. https://doi.org/10.1111/spc3.12341

Tetlock, P. E., Kristel, O. V., Elson, S. B., Green, M. C., & Lerner, J. S. (2000). The psychology of the unthinkable: Taboo trade-offs, forbidden base rates, and heretical counterfactuals. *Journal of Personality and Social Psychology*, *78*(5), 853–870. https://doi.org/10.1037//0022-3514.78.5.853

# Hierarchical Bayesian narrative-making under variable uncertainty

Alex Jinich-Diamant[a,b] and Leonardo Christov-Moore[b,c]

[a]Department of Anesthesiology, Department of Cognitive Science, University of California San Diego, La Jolla, CA 92093-0515, USA; [b]Multiscale Complexity

Institute, Centro Oaxaqueño de la Conciencia, 68287 Oaxaca, Mexico and
[c]Institute for Advanced Consciousness Studies, Santa Monica, CA 90403, USA.
aljinich@ucsd.edu
christovmoore@gmail.com
https://scholar.google.com/citations?user=PHJcx1IAAAAJ&hl=en

doi:10.1017/S0140525X22002643, e97

**Abstract**

While Conviction Narrative Theory correctly criticizes utility-based accounts of decision-making, it unfairly reduces probabilistic models to point estimates and treats affect and narrative as mechanistically opaque yet explanatorily sufficient modules. Hierarchically nested Bayesian accounts offer a mechanistically explicit and parsimonious alternative incorporating affect into a single biologically plausible precision-weighted mechanism that tunes decision-making toward narrative versus sensory dependence under varying uncertainty levels.

The authors argue that narrative is a mechanism by which decisions under radical uncertainty are made tractable to the predicting mind, a valuable criticism of reductionistic decision-making theories based on *homo economicus*. Conviction Narrative Theory (CNT) relies on affect and heuristics ("a fallible but useful system") as non-probabilistic mechanisms driving decision-making while simultaneously characterizing them as self-contained modules without a mechanistically explicit and biologically plausible characterization. In CNT, the brain operates in a fundamentally different way under risk than under radical uncertainty, explicitly assigning probabilities in the former and affectively responding to imagined futures in the latter. We believe the dichotomy between probabilistically explicit and radically uncertain decision-making modes to be inaccurate. Their claim assumes that the brain *does* explicitly assign probability point estimates to discrete outcomes under risk and that probabilities and narratives are mechanistic substitutes. However, predictive accounts of decision-making argue that the brain's computations are *well approximated* by Bayesian probability distributions, not that explicit single point estimates are actually assigned as in expected utility theory (Friston, 2010).

Hierarchical Bayesian accounts treat perception as a set of predictions constructed from (and bounded by) past experience, competing over which one best accounts for sensory evidence by minimizing prediction errors in a hierarchically nested fashion, from low-level sensory data (e.g., shiny objects with sharp edges) up to the causal narrative (e.g., glass broken by dog's wagging tail) (Friston, 2010). This does not require explicit probabilistic point estimates nor potentially infinite predictions. The hierarchical nature of perceptual processing means that the probabilistic process of choosing the best-fitting narrative is itself guided by the lower-level probabilistic processes constructing the sense data which the narratives compete to fit. Here, radical uncertainty simply means that more confidence is placed on already existing priors from higher up in the predictive hierarchy than on available sensory evidence when constructing the causal narrative that guides action. That higher-level predictions are more likely to be causal (e.g., dog broke glass) than lower-level sensory feature-based ones (e.g., furry tail next to sharp shards) accounts for why it may appear that we rely on narratives when uncertainty is higher. Explicit assignment of probability point estimates need not occur anywhere in this process, nor are heuristics and probabilities distinct operational modes. Rather, the same predictive system may be tuned to be more sensory versus narrative driven along a continuum between low and high uncertainty via precision-weighing of priors and evidence biologically implemented through neuromodulatory gain control (Barrett & Simmons, 2015). This account is a more parsimonious, mechanistically explicit, and biologically plausible account of decision-making under variable uncertainty.

The authors claim that "in lieu of probabilities to assess narratives, heuristics are used; […] instead of probabilities assigned to imagined futures, the single likeliest imagined future is adopted and evaluated. Rather than utilities assigned to particular outcomes over many dimensions, emotions are felt in response to imagined futures." Yet they fail to mechanistically characterize the heuristics and to say how they operate if not probabilistically. Doesn't "single likeliest imagined future" imply a probabilistic assessment? CNT tries to solve the cognitively intractable problem of evaluating simulated futures by relegating it to affect. Saying that "emotions are felt" fallaciously takes affect to be an opaque module existing safely outside of a probabilistic logic. In fact, we believe affect *is* what the authors are trying to account for in the first place – a fallible but useful shortcut (or heuristic) that probabilistically evaluates competing future simulations structured as culturally supplied narratives by fitting them to interoceptive sensory signals (Gendron, Mesquita, & Barrett, 2020). Theories of affect as homeostasis-driven prediction (Damasio & Carvalho, 2013) and as allostatic regulation (Kleckner et al., 2017) explain how probabilistic evaluation of interoceptive sensoria under varying uncertainty gives rise to feelings. These mechanistically explicit theories of affective processing treat feelings as low-dimensional representations of interoceptive predictions about the expected consequences of visceromotor commands on the body's internal milieu. Feelings thus encode the expected energy cost of an action that predictively responds to the organism's current situation via the same predictive hierarchy mentioned above, with narrative-like predictions at the top fitting sensory predictions at the bottom. CNT proposes that "explanatory fit is experienced affectively," but affective states can themselves drive narrative and narrative choice via interoception. In other words, their model is limited to the use of internal representations to account for exteroceptive data and misses interoception as the key element in action selection under uncertainty. If we take narrative to be the structure of higher-level priors in the nested predictive hierarchy, the gain control mechanisms that regulate the extent to which the affective system is sensory (bottom-up) driven versus prediction (top-down) driven accounts for why sometimes narrative is constructed based on evidence and why sometimes perceptual evidence is altered to fit an existing narrative, a fundamental phenomenon not explained by CNT.

While CNT is useful in that it offers a place for narrative and affect in probabilistic decision-making, it is less mechanistically explicit and parsimonious than hierarchical Bayesian accounts of affect, in which affect and narrative emerge from a single biologically plausible mechanism that tunes the system toward narrative versus sensory dependence as uncertainty increases. We believe the latter to better account for the role of narrative in action selection over CNT as the alternative to utility-based reductions of the human brain to *homo economicus*.

**Acknowledgements.** The authors thank Charlie C. M., Colla C. M., Sukka C. M., and Hoyden C. M.

**Financial support.** This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

**Competing interest.** None.

## References

Barrett, L. F., & Simmons, W. K. (2015). Interoceptive predictions in the brain. *Nature Reviews Neuroscience*, *16*(7), 419–429. https://doi.org/10.1038/nrn3950

Damasio, A., & Carvalho, G. B. (2013). The nature of feelings: Evolutionary and neurobiological origins. *Nature Reviews Neuroscience*, *14*(2), 143–152. https://doi.org/10.1038/nrn3403

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, *11*(2), 127–138. https://doi.org/10.1038/nrn2787

Gendron, M., Mesquita, B., & Barrett, L. F. (2020). The brain as a cultural artifact. Concepts, actions, and experiences within the human affective niche. In *Culture, mind, and brain: Emerging concepts, models, and applications* (pp. 188–222). Cambridge University Press. Retrieved from https://lirias.kuleuven.be/retrieve/596303

Kleckner, I. R., Zhang, J., Touroutoglou, A., Chanes, L., Xia, C., Simmons, W. K., … Feldman Barrett, L. (2017). Evidence for a large-scale brain system supporting allostasis and interoception in humans. *Nature Human Behaviour*, *1*(5), 1–14. https://doi.org/10.1038/s41562-017-0069

# Narrative as cultural attractor

James Holland Jones[a] and Calder Hilde-Jones[b]

[a]Department of Earth System Science, Stanford University, University of California San Diego, Stanford, CA 94305-4216 and [b]Department of Cognitive Science, University of California, San Diego, La Jolla, CA 92093-0515.
jhj1@stanford.edu
childejo@ucsd.edu
https://heeh.stanford.edu

doi:10.1017/S0140525X22002667, e98

**Abstract**

By structuring information in a systematic relational framework, narratives are cultural attractors that are particularly well-suited for transmission. The relational structure of narrative is partly what communicates causality, but this structure also complicates both transmission and selection on cultural elements by introducing correlations among narrative elements and between different narratives. These correlations have implications for adaptation, complexity, and robustness.

A key adaptive challenge humans face is the ability to make decisions under uncertainty, where learning is at best incomplete because of inherent complexity relative to information-processing capacities, an insufficient time span for learning, or nonstationarity of variability (Turner, Smaldino, Moya, & Jones, 2022). The general principle for the management of uncertainty is robustness (Kay & King, 2020): Solutions cannot be optimized for specific conditions, but need to be good enough for a wide range of conditions. The time frame for the emergence of the genus *Homo* coincided with a particularly variable period in Earth's history in a region with remarkably high-frequency variability (Antón, Potts, & Aiello, 2014). Moreover, this variability was nonstationary (Levin, 2015). Learning fully about environmental conditions was largely not possible because of this nonstationary, high-amplitude, fine-grain variability. This suggests that adaptive decision-making under uncertainty is a fundamental aspect of being human. The development of culture is clearly an essential feature of the success of the genus *Homo*, and *H. sapiens* in particular. Johnson et al.'s identification of narrative as a tool for dealing with radical uncertainty is a major conceptual contribution and has big implications for debates in the field of cultural evolution.

As it is shared knowledge, culture requires transmission. What distinguishes culture as an adaptive strategy is that it is shared and its advantages stem from relaxing the need to learn decisions anew across individuals and through time. What Sterelny (2017) calls the "California" school of cultural evolution emphasizes high-fidelity transmission of cultural elements effectively through copying with errors (e.g., Boyd & Richerson, 1985; Cavalli-Sforza & Feldman, 1981).

However, the organ that permits the development of culture, the brain, is not a passive copying machine; rather it is an information processor. In a world of infinite distractibility, elements of cultural knowledge must be attended to and meanings must be ascribed. As a result, it is likely that people transform elements of culture as they learn, process, and transmit them. This is the fundamental argument of Sperber (1996) and subsequent work in the tradition of cultural attractor theory, what Sterelny (2017) calls the "Parisian" school. The focus of the Parisians is on cultural attractors, which attempt to account for the fidelity and persistence of cultural ideas without the imitation of experts by novices that the California school presumes. While the idea that learners actively transform cultural information is appealing, the Parisian approach remains weak on formal mechanisms of transmission and is missing the strong modeling tradition of the California school (Sterelny, 2001, 2017). Without this, cultural attractor models lack explanatory power. Sterelny (2001, p. 848) suggests that the Parisian models "represent the consequences of transmission biases but they explain nothing about the sources of those biases." Narrative provides this.

While ideas may not replicate, we suggest that stories do. This makes narrative a very important type of cultural attractor. Narrative transforms ideas and generates emergent qualities of meaning, causal explanation, etc., but narratives are also uniquely transmittable. Stories are highly transmissible, even in the presence of substantial noise or long transmission chains (Mesoudi, Whiten, & Dunbar, 2006). By facilitating transmission, narrative also functions as a means to distribute cognition beyond the single individual. Narrative improves recall and interest in information, aiding that crucial limiting resource, attention (Glaser, Garsoffky, & Schwan, 2009). By facilitating attention, retention, and transmission, narrative helps accomplish the crucial elements of culture, namely, extrasomatic information storage and processing.

Narrative includes a number of features that make it useful for encoding and transmitting information under conditions of uncertainty. Narrative provides (1) a mechanism for transmitting complex information that goes well beyond imitation, (2) causal structure, thereby facilitating exploration, and (3) coherence, which ensures robustness. Currie and Sterelny (2017, p. 19) note, "As our information about the causal background is enriched, coherence becomes an increasingly important, increasingly demanding constraint." Checks on coherence ensure a degree of fidelity in transmission and provide the foundation for selection to work.

As Johnson et al. note, narratives are causal explanations. Their causal structure, representable as a directed graph (as in their Fig. 3), means that the elements of narratives are *relational*. Through their relationship, as indicated by an arc in the directed

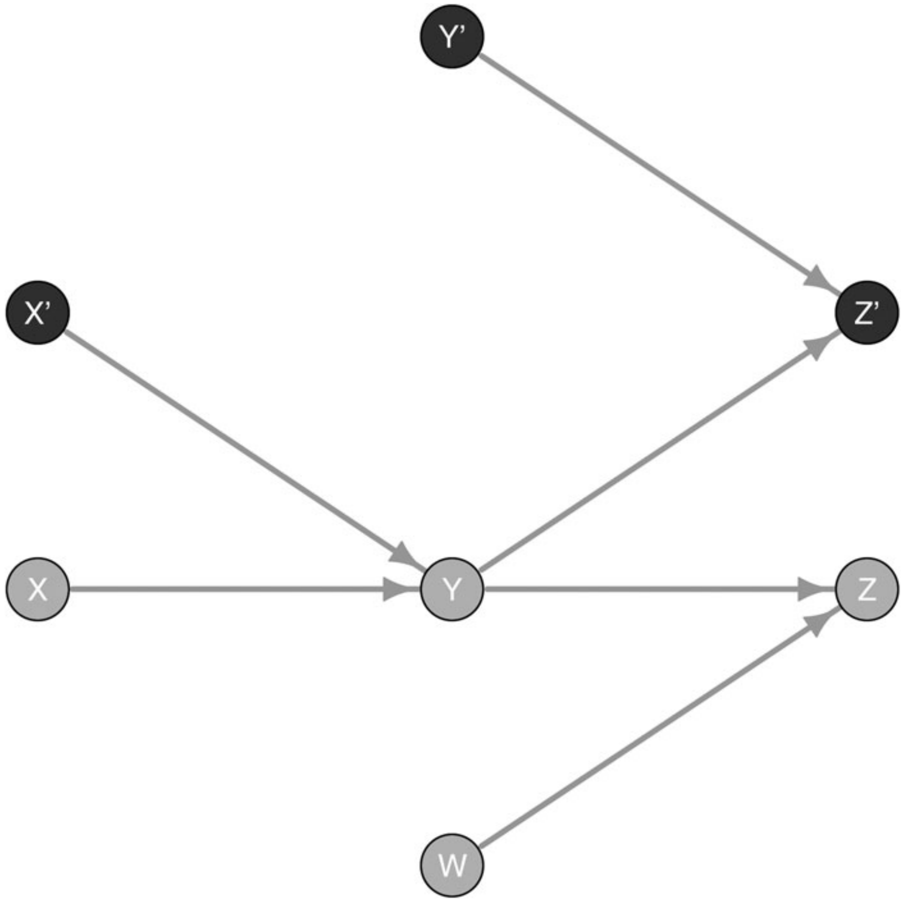


**Figure 1** (Jones and Hilde-Jones) Two simple narratives, represented as directed graphs, that provide a causal structure for two outcomes (Z and Z') with a shared element, Y.

graph, two elements cease to be independent of each other. Their transmission, and selection on them, will be correlated. Correlation greatly increases the complexity of both population dynamics (e.g., Tuljapurkar, 1990) and the response to selection (e.g., Lande, 1979). Very little work in cultural evolution has addressed the complications of indirect selection, but some initial forays have been made (Yeh, Fogarty, & Kandler, 2019).

Narratives, as transmissible cultural attractors, complicate selection on cultural elements. In addition to the relational nature of the elements of a narrative, highly compelling narratives attract associated concepts through semantic association, metaphor, and connotation (as noted in sect. 5.1.3). If we want to understand the dynamics of selection on particular cultural traits, we need to consider them as multivariate traits because of the emergence of covariance between elements both within the narrative and between the narrative itself and associated narratives. This covariance means that cultural traits will respond to both direct selection, as with standard cultural-evolutionary models, but also to indirect selection on covarying cultural traits. Figure 1 provides a demonstration of two simple narratives that provide a causal structure for two outcomes (Z and Z'), with common element across the two narratives of Y. Not only will the elements within a narrative be correlated from the perspective of selection acting on their frequency, the two narratives themselves will be correlated.

This has many implications. One important implication, particularly for the early selective advantage of narrative transmission, is that we expect narrative structures to be more robust to mistakes about specific elements. Evolutionary robustness is the persistence of an organismal trait under perturbations (Wagner, 2013). Of course, narratives as cultural attractors can shape meaning in potentially harmful ways as well. A high-coherence narrative can capture otherwise orthogonal information or can be resistant to updating as contradictory evidence accumulates.

**Financial support.** This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

**Competing interest.** None.

## References

Antón, S. C., Potts, R., & Aiello, L. C. (2014). Evolution of early *Homo*: An integrated biological perspective. *Science*, *345*(6192), 1236828. doi: 10.1126/science.1236828

Boyd, R., & Richerson, P. J. (1985). *Culture and the evolutionary process*. University of Chicago Press.

Cavalli-Sforza, L. L., & Feldman, M. W. (1981). *Cultural transmission and evolution: A quantitative approach*. Princeton University Press.

Currie, A., & Sterelny, K. (2017). In defence of story-telling. *Studies in History and Philosophy of Science Part A*, *62*, 14–21. doi: 10.1016/j.shpsa.2017.03.003

Glaser, M., Garsoffky, B., & Schwan, S. (2009). Narrative-based learning: Possible benefits and problems. *Communications*, *34*(4), 429. doi: 10.1515/COMM.2009.026

Kay, J. A., & King M. A. (2020). *Radical uncertainty: Decision-making beyond the numbers*. Norton.

Lande, R. (1979). Quantitative genetic analysis of multivariate evolution, applied to brain: Body size allometry. *Evolution*, *33*(1), 402–416. doi:10.1111/j.1558-5646.1979.tb04694.x

Levin, N. E. (2015). Environment and climate of early human evolution. *Annual Review of Earth and Planetary Sciences*, *43*(1), 405–429. doi: 10.1146/annurev-earth-060614-105310

Mesoudi, A., Whiten, A., & Dunbar, R. (2006). A bias for social information in human cultural transmission. *British Journal of Psychology*, *97*(3), 405–423. doi: 10.1348/000712605X85871

Sperber, D. (1996). *Explaining culture: A naturalistic approach*. Blackwell.

Sterelny, K. (2001). [Review of explaining culture: A naturalistic approach, by D. Sperber]. *Mind*, *110*(439), 845–854. http://www.jstor.org/stable/3093692
Sterelny, K. (2017). Cultural evolution in California and Paris. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *62*, 42–50. doi: 10.1016/j.shpsc.2016.12.005
Tuljapurkar, S. (1990). *Population dynamics in variable environments*. Lecture Notes in Biomathematics, vol. *85*. Springer-Veralg.
Turner, M. A., Smaldino, P., Moya, C., & Jones, J. H. (2022). Some forms of uncertainty may suppress the evolution of social learning. *PsyArXiv*, 1 July 2022, 1–7. doi: 10.31234/osf.io/dzteh
Wagner, A. (2013). *Robustness and evolvability in living systems*. Princeton University Press.
Yeh, D. J., Fogarty, L., & Kandler, A. (2019). Cultural linkage: The influence of package transmission on cultural dynamics. *Proceedings of the Royal Society B: Biological Sciences*, *286*(1916), 20191951. doi: 10.1098/rspb.2019.1951

# Do conviction narratives drive individual decisions?

Mervyn King

London School of Economics, House of Lords, London SW1A 0PW, UK.
office@mervynking.com

doi:10.1017/S0140525X22002539, e99

**Abstract**

Conviction Narrative Theory is an interesting and plausible way to think about how individuals make decisions when quantitative assessments are not open to them. The question I pose is the following. Is there any general result about how decisions should be taken independent of the particular circumstances of that decision?

Decision-making under uncertainty has tended to proceed on the assumption that there is a "right" and a "wrong" way of making decisions. There is a large literature on the assumptions behind the view that individuals or organisations should maximise a measure of expected "utility" but all approaches along these lines require the assumption that there are subjective probabilities attached to every conceivable state of the world. In a world of radical uncertainty, however, no such probabilities exist (Kay & King, 2020). How then should individuals make decisions in such a world?

There are two immediate implications of radical uncertainty. First, optimising behaviour cannot be defined. Second, there is no single rule of thumb that can replace optimising behaviour: Individuals must learn to cope with circumstances as they arise. What modes of thought are useful in confronting the situation of this kind? Johnson et al. recognise that when quantitative assessments of uncertainty are impossible, individuals try to understand radical uncertainty in terms of narratives. This is how people communicate with each other. But they go much further in analysing how narratives are constructed and communicated to others. Their Conviction Narrative Theory (CNT) argues that "narratives arise from the interplay between individual cognition and the social environment, with reasoners adopting a narrative that feels 'right' to explain the available data; using that narrative to imagine plausible futures; and affectively evaluating those imagined futures to make a choice." This is an interesting and plausible way to think about how individuals make decisions when quantitative assessments are not open to them. As the authors comment, "we use narratives to make sense of the past, imagine the future, commit to action, and share these judgments and choices with others."

Narratives are used to evaluate "imagined futures." In turn, decisions are made through the adoption of heuristics which are simple rules of thumb that precisely, because they are simple, are likely to be robust. All of this is compelling. The paper provides a rich framework within which to think about what constitutes rational decision-making in a world of radical uncertainty. But the question I would pose is the following. Although the authors start and continue with revealing examples of decisions, the conceptual apparatus of CNT is not used to provide answers except in a general and abstract form. Is it possible to do so, or are we driven to the proposition that only an analysis of specific circumstances can explain why a decision was taken? Is there any general result that we can discuss independent of the particular circumstances of that decision? The intellectual attractions of a general theory are clear. But is this an illusion in our present state of knowledge? More fundamentally, in a world of radical uncertainty can there be a general theory at all?

The authors are clearly conscious of this issue. They argue, correctly in my view, that they are tackling a "grand" problem. But the question of how far we can get in constructing the general theory remains open. Nevertheless, as the authors conclude, "the preoccupation of much decision-making research with optimality – whether in assumption or subversion – might profitably yield some ground to the more basic question of how, under radical uncertainty and fuzzy evaluation, we gain conviction to act at all."

Social scientists across a variety of disciplines are needed to make progress in this important area.

**Financial support.** This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

**Competing interest.** None.

## Reference

Kay, J. A. & King, M. A. (2020). *Radical uncertainty: Decision-making for an unknowable future*. The Bridge Street Press.

# Adaptive narratives and fantastical falsehoods?

Aaron D. Lightner

Department of the Study of Religion, Aarhus University, Jens Chr. Skous Vej 3, 8000 Aarhus C, Denmark.
adlightner@cas.au.dk

doi:10.1017/S0140525X22002588, e100

**Abstract**

Johnson et al. make a strong case for Conviction Narrative Theory, but it remains unclear why so many adaptive narratives include supernatural causes and other falsehoods. Focusing on religions, I argue that an adaptive decision-making system might include supernatural falsehoods because they simplify complex problems, they are sensitive to long-term incentives, and they evoke strong emotions in a communicative context.

Conviction Narrative Theory (CNT) resonates with the anthropological evidence that societies use *narratives* to frame uncertain decisions (Douglas, 2013). People evaluate narratives with a blend of emotion and intuition (Zinn, 2016), and CNT convincingly outlines how this might be part of an adaptive decision-making system under radical uncertainty. And yet anthropological evidence also shows that high-risk and uncertain decisions frequently involve *religious* narratives with supernatural causes and other falsehoods (Evans-Pritchard, 1937; Malinowski, 1932). If our narratives are adaptive for dealing with uncertain fitness-relevant problems, then why are so many narratives false?

A popular answer is that our evolved cognition is like a flawed superpower: Humans might be generally sensitive to the statistical and causal structure of their environments (Sperber, Premack, & Premack, 1995), but they are also prone to biases and misleading intuitions that disrupt these capabilities (Kahneman, 2013). An alternative answer, which I advocate here, is that religious falsehoods *are* part of a broader tendency to generate useful, ecologically rational narratives (Lightner & Hagen, 2022). Under radical uncertainty, where information is scarce and the data-generating processes are unknown, how *should* people think about fitness-relevant challenges like natural hazards or social conflicts? I will discuss three ingredients that supernatural falsehoods might add to decision-making narratives.

The first ingredient is that they *simplify complex problems*. Contra Johnson et al. in section 6.1, there are strong theoretical reasons to suggest that simple narratives explain *more* of the data under radical uncertainty, because they are less sensitive to variance around future observations (i.e., they improve a bias-variance tradeoff; Brighton & Gigerenzer, 2012). When explaining complex phenomena, a simpler narrative that ignores the idiosyncratic details of individual observations is more likely to generalize to other scenarios with different background circumstances (Quillien, 2020; Woodward, 2006).

A familiar example is the intentional stance (Dennett, 1987). Humans can predict and navigate social behavior because we ascribe simple, reductive mental state concepts to a truly complex and mysterious type of data-generator: other minds (Gerstenberg & Tenenbaum, 2017). Religious falsehoods similarly tend to impute oversimplified, anthropomorphic falsehoods onto complex phenomena. For example, many societies depend on forest resources. Forests are complex ecological systems (Filotas et al., 2014), where many species interact in a dynamic "web-like structure" resembling social causation (target article, sect. 6.1). Animistic religions often construe these forest systems as having person-like entities that purposefully engage in helpful, antagonistic, and communicative relationships with each other (Atran et al., 2002; Ingold, 2006). This perspective might be dismissed as superstitious, but it can serve the genuinely useful and otherwise difficult task of making sense of a complex, dynamic ecological network (ojalehto mays, Seligman, & Medin, 2020).

The second ingredient is that supernatural falsehoods are often *sensitive to long-term outcomes that matter*. Consider, for example, the Australian Martu landscape burning practices that ultimately lead to continuing food availability (Bird, Tayor, Codding, & Bird, 2013). These practices create a collective action problem because the payoffs are only high if people participate in the required labor (Lightner & Purzycki, 2023). Rather than resorting to decision or game theory, Martu view their burning practices as part of a sacred narrative where failing to participate would bring about the end of the world (Bird, Bird, Codding, & Taylor, 2016). In a real sense, this has useful causal information: In their harsh and arid environment, a collapse of the food supply *would* be a dire outcome.

Even explanations for "randomness" vary across cultures in ways that seem sensitive to long-term ecological risks (Tucker, 2017). Howell (2012), for example, observed that the Malaysian Chewong explain misfortune in terms of spiritual retribution for norm violations rather than chance. In a sense, this seems irrational because it posits a tapestry of non-human agents who sanction inauspicious practices. In another sense, Howell argues, there's an ecologically relevant logic to their view of randomness, because Chewong social rules "prevent risk situations from occurring" while recognizing the fact that "in performing their daily tasks, humans are not separate from the natural world" (p. 9).

Interestingly, Johnson et al. discuss an unpublished experiment that also seems to suggest that people think of randomness as an explanation, rather than a lack thereof (Johnson, Matiashvili, & Tuckett, 2019; sect. 7.1.1). They also mention that participants tended to think that unexplained price changes were signals when positive. If participants had a long enough time horizon in mind, then this does not seem like a particularly irrational heuristic because major stock market indices do tend to increase over a long enough timeframe (Kim & Ryoo, 2011). Like the Martu and Chewong, when given a noisy process with scarce information, participants might have been sensitive to the feedback they have received in the long run.

The third ingredient is that supernatural falsehoods can effectively manipulate behavior by *evoking strong emotions when they are communicated*. Consider the many sources of danger that forests can contain at night (e.g., predators, enemies, falling hazards). Rather than enumerating and weighing all the possibilities, many societies use narratives to evoke salient fears that effectively keep others from harm's way (e.g., "avoid the forest at night because monsters lurk in there") (Morin & Sobchuk, 2022; Sugiyama, Sugiyama, Slingerland, & Collard, 2011). Supernatural punishment is an especially common and effective tool for motivating cooperative behavior because it presents decision-makers with a distressing imagined future for breaking the rules (Bendixen & Purzycki, 2017; Fitouchi & Singh, 2021). Indeed, supernatural appeals across cultures are not about random subject matter; they are usually about locally important socioecological challenges like resource management and cooperative conflicts (Bendixen et al., 2021; Purzycki, Bendixen, Lightner, & Sosis, 2022). They are also sensitive to long-term causation: Cheating might pay in the short-term, but losing a cooperative partner can genuinely present a long-term opportunity cost (Delton, Krasnow, Cosmides, & Tooby, 2011).

Finding the narrative that "feels right," as CNT puts it, can help bypass a considerable amount of noise when responding to a complex and mysterious problem. More importantly, religious narratives, though false or seemingly irrational, often turn out to reflect concerns about behaviors that matter for people's livelihoods. With radical uncertainty, they are often all we have.

**Acknowledgement.** ADL thanks Benjamin Purzycki for feedback.

**Financial support.** ADL is supported by the Aarhus University Research Foundation.

**Competing interest.** None.

## References

Atran, S., Medin, D., Ross, N., Lynch, E., Vapnarsky, V., Ek, E., … Baran, M. (2002). Folkecology, cultural epidemiology, and the spirit of the commons: A garden experiment in the Maya lowlands, 19912001. *Current Anthropology*, *43*(3), 421–450.

Bendixen, T., Apicella, C., Atkinson, Q., Cohen, E., Henrich, J., McNamara, R. A., … Purzycki, B. G. (2021). Appealing to the minds of gods: A novel cultural evolutionary account of religious appeals and an empirical assessment using ethnographic data from eight diverse societies.

Bendixen, T., & Purzycki, B. G. (2017). Peering into the minds of gods: What cross-cultural variation in gods' concerns can tell us about the evolution of religion. *Journal for the Cognitive Science of Religion*, *5*(2), 142–165.

Bird, D. W., Bird, R. B., Codding, B. F., & Taylor, N. (2016). A landscape architecture of fire: Cultural emergence and ecological pyrodiversity in Australia's western desert. *Current Anthropology*, *57*(S13), S65–S79.

Bird, R. B., Tayor, N., Codding, B. F., & Bird, D. W. (2013). Niche construction and dreaming logic: Aboriginal patch mosaic burning and varanid lizards (*Varanus gouldii*) in Australia. *Proceedings of the Royal Society B: Biological Sciences*, *280*(1772), 20132297.

Brighton, H., & Gigerenzer, G. (2012). Homo heuristicus and the bias-variance dilemma. In J. Schulkin (Ed.), *Action, perception and the brain* (pp. 68–91). Springer.

Delton, A. W., Krasnow, M. M., Cosmides, L., & Tooby, J. (2011). Evolution of direct reciprocity under uncertainty can explain human generosity in one-shot encounters. *Proceedings of the National Academy of Sciences*, *108*(32), 13335–13340.

Dennett, D. C. (1987). *The intentional stance*. MIT Press.

Douglas, M. (2013). *Risk and blame: Essays in cultural theory*. Routledge.

Evans-Pritchard, E. E. (1937). *Witchcraft, oracles, and magic among the Azande*. Clarendon Press, abridged with an introd. by eva gillies edition.

Filotas, E., Parrott, L., Burton, P. J., Chazdon, R. L., Coates, K. D., Coll, L., … Messier, C. (2014). Viewing forests through the lens of complex systems science. *Ecosphere*, *5*(1), 1–23.

Fitouchi, L., & Singh, M. (2021). Supernatural punishment beliefs as cognitively compelling tools of social control. *Current Opinion in Psychology*, *44*, 252–257.

Gerstenberg, T., & Tenenbaum, J. B. (2017). Intuitive theories. *Oxford handbook of causal reasoning* (pp. 515–548). Oxford University Press.

Howell, S. (2012). Knowledge, morality, and causality in a "luckless" society: The case of the Chewong in the Malaysian rain forest. *Social Analysis*, *56*(1), 133–147.

Ingold, T. (2006). Rethinking the animate, re-animating thought. *Ethnos*, *71*(1), 9–20.

Johnson, S. G. B., Matiashvili, T., & Tuckett, D. (2019). Explaining the past, predicting the future: How attributions for past price changes affect price expectations. Manuscript in preparation.

Kahneman, D. (2013). *Thinking, fast and slow*. Farrar, Straus and Giroux, 1st pbk. ed edition.

Kim, J. H., & Ryoo, H. H. (2011). Common stocks as a hedge against inflation: Evidence from century-long us data. *Economics Letters*, *113*(2), 168–171.

Lightner, A. D., & Hagen, E. (2022). All models are wrong, and some are religious: Supernatural explanations as abstract and useful falsehoods about complex realities. *Human Nature*, *33*, 425–462.

Lightner, A. D., & Purzycki, B. G. (2023). Game theoretical aspects of the minds of gods. In B. G. Purzycki & T. Bendixen (Eds.), *The minds of gods: A cross-disciplinary survey* (pp. 133–147). Bloomsbury Publishing.

Malinowski, B. (1932). *Argonauts of the Western Pacific*. George Routledge and Sons, Limited.

Morin, O., & Sobchuk, O. (2022). Why monsters are dangerous.

ojalehto mays, B., Seligman, R., & Medin, D. L. (2020). Cognition beyond the human: Cognitive psychology and the new animism. *Ethos*, *48*(1), 50–73.

Purzycki, B. G., Bendixen, T., Lightner, A. D., & Sosis, R. (2022). Gods, games, and the socioecological landscape. *Current Research in Ecological and Social Psychology*, *3*, 100057.

Quillien, T. (2020). When do we think that x caused y? *Cognition*, *205*, 104410.

Sperber, D., Premack, D., & Premack, A. J. (1995). *Causal cognition: A multidisciplinary debate*. Clarendon Press Oxford.

Sugiyama, M. S., & Sugiyama, L. S. (2011). Once the child is lost he dies: Monster stories vis-a-vis the problem of errant children. In E. Slingerland & M. Collard (Eds.), *Creating consilience: Integrating the sciences and the humanities* (pp. 351–371). Oxford Academic Press.

Tucker, B. (2017). From risk and time preferences to cultural models of causality: On the challenges and possibilities of field experiments, with examples from rural southwestern Madagascar. In Jeffrey R. Stevens (Ed.), *Impulsivity: How time and risk influence decision making* (pp. 61–114). Springer.

Woodward, J. (2006). Sensitive and insensitive causation. *The Philosophical Review*, *115*(1), 1–50.

Zinn, J. O. (2016). "In-between" and other reasonable ways to deal with risk and uncertainty: A review article. *Health, Risk & Society*, *18*(7–8), 348–366.


# When radical uncertainty is too much: Clinical aspects of Conviction Narrative Theory


Omer Linkovski[a] and Renana Eitan[b,c]

[a]Department of Psychology and The Gonda Interdisciplinary Brain Research Center, Bar-Ilan University, Ramat-Gan 5290002, Israel; [b]Department of Psychiatry, Tel-Aviv Medical Center, Tel-Aviv 6423906, Israel and [c]Department of Psychiatry, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115, USA.

Omer.linkovski@biu.ac.il
renanae@tlvmc.gov.il
https://linkovskilab.com
https://www.tasmc.org.il/sites/en/Personnel/pages/eitan-renana.aspx


doi:10.1017/S0140525X2200259X, e101


### Abstract

We propose extrapolating Conviction Narrative Theory (CNT) to clinical psychology and psychiatry. We demonstrate how CNT principles may benefit assessment, therapy, and possibly even modify public health views of neuropsychiatric disorders. Our commentary focuses on hoarding disorder as a model, elaborates on discrepancies in the scientific literature and suggests how the CNT may resolve them.


Johnson et al. provide a thought-provoking and insightful synthesis of ideas advancing our understanding of decision-making. We suggest that the reach of the Conviction Narrative Theory (CNT) affects far broader psychology fields than Johnson et al. propose. We argue for widening CNT's scope to neuropsychiatric disorders, psychotherapy, and their impact on public health.

Considering the CNT as it applies to generic decision-making under radical uncertainty may limit its impact. Difficulties tolerating uncertainty and resolving it are central in multiple neuropsychiatric disorders (Gillett, Bilek, Hanna, & Fitzgerald, 2018; Lebert, Turkington, Freeston, & Dudley, 2021). Individuals often portray such difficulties using subjective reports, objective measures, or both. There are discrepancies between subjective reports and objective tasks assessing resolution of uncertainty or decision-making under uncertainty. Healthy adults' objective resolution of uncertainty may differ under certain conditions, but their subjective estimate of uncertainty remains constant in these different conditions (Linkovski, Rodriguez, Wheaton, Henik, & Anholt, 2021). In other studies, the association between subjective intolerance of uncertainty (IU) and subjective anxiety severity was higher than the association between IU and objective decision-making under uncertainty (Luhmann, Ishida, & Hajcak, 2011). The CNT accounts for these discrepancies as the objective uncertainty measures include a finite set of options and the subjective measures assess IU in real life, where radical uncertainty is abundant.

The discrepancy between objective and subjective uncertainty measures is imperative for neuropsychiatric disorders. Difficulties resolving uncertainty are transdiagnostic and may serve as a treatment target but are often measured with subjective assessments (McEvoy, Hyett, Shihata, Price, & Strachan, 2019).

Multiple clinical cohorts report increased subjective IU (Gillett et al., 2018) and subjective IU predicts response to psychotherapy (Castriotta, Dozier, Taylor, Mayes, & Ayers, 2019). Objective uncertainty measures most often include monetary decision-making or cognitive reasoning (Aranovich, Cavagnaro, Pitt, Myung, & Mathews, 2017; Pushkarskaya et al., 2015, 2017; Ruderman et al., 2016; Strauss et al., 2020; Zald & Treadway, 2017) as they are derived from rational approaches and are easily inferred. These tasks informed our understanding of neuropsychiatric disorders and hold some predictive validity to developing psychopathologies (Ruderman et al., 2016). Yet there are discrepancies between these objective tasks and subjective assessments. For example, hoarding disorder (HD) patients struggle with letting go of items, irrespective of their objective value (American Psychiatric Association, 2013). These patients have elevated subjective IU scores and often report struggling with uncertainty in their daily life (Wheaton, Abramowitz, Jacoby, Zwerling, & Rodriguez, 2016). Objective decision-making tasks suggest that HD patients have intact decision-making under uncertainty, although they may be less flexible (Pushkarskaya et al., 2017) and that these patients are less sensitive to losses than healthy adults (Aranovich et al., 2017). In real life however, HD patients may be extremely sensitive to losing items (Orr, Preston-Shoot, & Braye, 2019). This discrepancy may be reduced if we incorporate CNT principles in objective tasks. For example, using emotionally laden tasks yielded significant leaps in our understanding of neuronal mechanisms of HD (Tolin et al., 2012). We hypothesize that experimental tasks using multiple decision options, affect-laden stimuli, and having individuals explain their narrative fragments or engage in affective valuation or simulation of decision outcomes will result in much greater differences between HD and non-HD cohorts.

The CNT framework aligns with psychotherapeutic interventions. Working with clients to expose, expand, and alter their narratives is a common goal of many psychotherapies (Goldblatt, Briggs, Lindner, Schechter, & Ronningstam, 2015; Rhodes, 2013). In HD, first-line psychotherapies aim to modify thoughts, beliefs, and decision-making processes assigned to items and to the self (Mathews et al., 2018; Steketee & Frost, 2013; Tolin, Frost, & Steketee, 2007). The CNT may define these thoughts, beliefs, and decision-making processes as narratives or narrative fragments. In line with CNT principles, current treatments have therapists or volunteers aiding HD clients assess their narratives of items and verbally simulate outcomes of their decisions in their homes – such practices may increase affective reasoning and widen the narrative fragments (Crone, Angel, Isemann, & Norberg, 2020; Linkovski et al., 2018; Muroff & Otte, 2019). The CNT suggests that narratives bind processes and perception to guide actions (target article). Therapists may work with clients on explicitly communicating new narrative fragments, which in turn will modify clients' narrative and alter their actions in a more profound way.

CNT discusses shared narratives and their role in economic outcomes. Neuropsychiatric disorders are a leading cause in reducing economic productivity (Chen, Kuhn, Prettner, & Bloom, 2018; James et al., 2018; Wittchen et al., 2011). Economic models suggest that major depressive disorder alone costs the US economy 326 billion dollars per year (Greenberg et al., 2021). These disorders are associated with societal stigma (Chasson, Guy, Bates, & Corrigan, 2018; Corrigan, 2004) which is internalized and affects treatment-seeking and informing mental health status with employers (Clement et al., 2015; Schnyder, Panczak, Groth, & Schultze-Lutter, 2017). Increasing help-seeking behaviors may lessen the economic burden of neuropsychiatric disorders. Extrapolating the CNT suggests that changing the shared narrative of mental health may reduce the economic costs of neuropsychiatric disorders.

To summarize, we propose that CNT can benefit how we study and treat neuropsychiatric disorders and help decision-makers and health systems lessen the global burden these disorders cause.

**Acknowledgements.** None.

**Financial support.** The current work was partially funded by an Israel Science Foundation Grant to Dr. Eitan (#2128/19).

**Competing interest.** None.

## References

American Psychiatric Association (Ed.). (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (5th ed). American Psychiatric Association.

Aranovich, G. J., Cavagnaro, D. R., Pitt, M. A., Myung, J. I., & Mathews, C. A. (2017). A model-based analysis of decision making under risk in obsessive-compulsive and hoarding disorders. *Journal of Psychiatric Research*, *90*, 126–132. https://doi.org/10.1016/j.jpsychires.2017.02.017

Castriotta, N., Dozier, M. E., Taylor, C. T., Mayes, T., & Ayers, C. R. (2019). Intolerance of uncertainty in hoarding disorder. *Journal of Obsessive-Compulsive and Related Disorders*, *21*, 97–101. https://doi.org/10.1016/j.jocrd.2018.11.005

Chasson, G. S., Guy, A. A., Bates, S., & Corrigan, P. W. (2018). They aren't like me, they are bad, and they are to blame: A theoretically-informed study of stigma of hoarding disorder and obsessive-compulsive disorder. *Journal of Obsessive-Compulsive and Related Disorders*, *16*, 56–65. https://doi.org/10.1016/j.jocrd.2017.12.006

Chen, S., Kuhn, M., Prettner, K., & Bloom, D. E. (2018). The macroeconomic burden of noncommunicable diseases in the United States: Estimates and projections. *PLoS ONE*, *13*(11), e0206702. https://doi.org/10.1371/journal.pone.0206702

Clement, S., Schauman, O., Graham, T., Maggioni, F., Evans-Lacko, S., Bezborodovs, N., … Thornicroft, G. (2015). What is the impact of mental health-related stigma on help-seeking? A systematic review of quantitative and qualitative studies. *Psychological Medicine*, *45*(1), 11–27. https://doi.org/10.1017/S0033291714000129

Corrigan, P. (2004). How stigma interferes with mental health care. *American Psychologist*, *59*(7), 614–625. https://doi.org/10.1037/0003-066X.59.7.614

Crone, C., Angel, Z., Isemann, S., & Norberg, M. M. (2020). Clutter-Buddies: A volunteer program to assist clients undergoing group cognitive behavioural therapy. *Journal of Obsessive-Compulsive and Related Disorders*, *27*, 100559. https://doi.org/10.1016/j.jocrd.2020.100559

Gillett, C. B., Bilek, E. L., Hanna, G. L., & Fitzgerald, K. D. (2018). Intolerance of uncertainty in youth with obsessive-compulsive disorder and generalized anxiety disorder: A transdiagnostic construct with implications for phenomenology and treatment. *Clinical Psychology Review*, *60*, 100–108. https://doi.org/10.1016/j.cpr.2018.01.007

Goldblatt, M. J., Briggs, S., Lindner, R., Schechter, M., & Ronningstam, E. (2015). Psychodynamic psychotherapy with suicidal adolescents. *Psychoanalytic Psychotherapy*, *29*(1), 20–37. https://doi.org/10.1080/02668734.2015.1004188

Greenberg, P. E., Fournier, A.-A., Sisitsky, T., Simes, M., Berman, R., Koenigsberg, S. H., & Kessler, R. C. (2021). The economic burden of adults with major depressive disorder in the United States (2010 and 2018). *PharmacoEconomics*, *39*(6), 653–665. https://doi.org/10.1007/s40273-021-01019-4

James, S. L., Abate, D., Abate, K. H., Abay, S. M., Abbafati, C., Abbasi, N., … Murray, C. J. L. (2018). Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: A systematic analysis for the Global Burden of Disease Study 2017. *The Lancet*, *392*(10159), 1789–1858. https://doi.org/10.1016/S0140-6736(18)32279-7

Lebert, L., Turkington, D., Freeston, M., & Dudley, R. (2021). Rumination, intolerance of uncertainty and paranoia in treatment resistant psychosis. *Psychosis*, *13*(1), 65–70. https://doi.org/10.1080/17522439.2020.1798489

Linkovski, O., Rodriguez, C. I., Wheaton, M. G., Henik, A., & Anholt, G. E. (2021). Momentary induction of inhibitory control and its effects on uncertainty. *Journal of Cognition*, *4*(1), 10. https://doi.org/10.5334/joc.133

Linkovski, O., Zwerling, J., Cordell, E., Sonnenfeld, D., Willis, H., La Lima, C. N., … Rodriguez, C. I. (2018). Augmenting buried in treasures with in-home uncluttering practice: Pilot study in hoarding disorder. *Journal of Psychiatric Research*, *107*, 145–150. https://doi.org/10.1016/j.jpsychires.2018.10.001

Luhmann, C. C., Ishida, K., & Hajcak, G. (2011). Intolerance of uncertainty and decisions about delayed, probabilistic rewards. *Behavior Therapy*, *42*(3), 378–386. https://doi.org/10.1016/j.beth.2010.09.002
Mathews, C. A., Mackin, R. S., Chou, C.-Y., Uhm, S. Y., Bain, L. D., Stark, S. J., … Delucchi, K. (2018). Randomised clinical trial of community-based peer-led and psychologist-led group treatment for hoarding disorder. *BJPsych Open*, *4*(4), 285–293. https://doi.org/10.1192/bjo.2018.30
McEvoy, P. M., Hyett, M. P., Shihata, S., Price, J. E., & Strachan, L. (2019). The impact of methodological and measurement factors on transdiagnostic associations with intolerance of uncertainty: A meta-analysis. *Clinical Psychology Review*, *73*, 101778. https://doi.org/10.1016/j.cpr.2019.101778
Muroff, J., & Otte, S. (2019). Innovations in CBT treatment for hoarding: Transcending office walls. *Journal of Obsessive-Compulsive and Related Disorders*, *23*, 100471. https://doi.org/10.1016/j.jocrd.2019.100471
Orr, D. M. R., Preston-Shoot, M., & Braye, S. (2019). Meaning in hoarding: Perspectives of people who hoard on clutter, culture and agency. *Anthropology & Medicine*, *26*(3), 263–279. https://doi.org/10.1080/13648470.2017.1391171
Pushkarskaya, H., Tolin, D., Ruderman, L., Henick, D., Kelly, J. M., Pittenger, C., & Levy, I. (2017). Value-based decision making under uncertainty in hoarding and obsessive-compulsive disorders. *Psychiatry Research*, *258*, 305–315. https://doi.org/10.1016/j.psychres.2017.08.058
Pushkarskaya, H., Tolin, D., Ruderman, L., Kirshenbaum, A., Kelly, J. M., Pittenger, C., & Levy, I. (2015). Decision-making under uncertainty in obsessive–compulsive disorder. *Journal of Psychiatric Research*, *69*, 166–173. https://doi.org/10.1016/j.jpsychires.2015.08.011
Rhodes, J. (2013). *Narrative CBT: Distinctive features*. Routledge.
Ruderman, L., Ehrlich, D. B., Roy, A., Pietrzak, R. H., Harpaz-Rotem, I., & Levy, I. (2016). Posttraumatic stress symptoms and aversion to ambiguous losses in combat veterans. *Depression and Anxiety*, *33*(7), 606–613. https://doi.org/10.1002/da.22494
Schnyder, N., Panczak, R., Groth, N., & Schultze-Lutter, F. (2017). Association between mental health-related stigma and active help-seeking: Systematic review and meta-analysis. *The British Journal of Psychiatry*, *210*(4), 261–268. https://doi.org/10.1192/bjp.bp.116.189464
Steketee, G., & Frost, R. O. (2013). *Treatment for hoarding disorder: workbook* (2nd ed.). Oxford University Press.
Strauss, A. Y., Fradkin, I., McNally, R. J., Linkovski, O., Anholt, G. E., & Huppert, J. D. (2020). Why check? A meta-analysis of checking in obsessive-compulsive disorder: Threat vs. distrust of senses. *Clinical Psychology Review*, *75*, 101807. https://doi.org/10.1016/j.cpr.2019.101807
Tolin, D. F., Frost, R. O., & Steketee, G. (2007). *Buried in treasures: Help for compulsive acquiring, saving, and hoarding* (1st ed.). Oxford University Press.
Tolin, D. F., Stevens, M. C., Villavicencio, A. L., Norberg, M. M., Calhoun, V. D., Frost, R. O., … Pearlson, G. D. (2012). Neural mechanisms of decision making in hoarding disorder. *Archives of General Psychiatry*, *69*(8), 832–841.
Wheaton, M. G., Abramowitz, J. S., Jacoby, R. J., Zwerling, J., & Rodriguez, C. I. (2016). An investigation of the role of intolerance of uncertainty in hoarding symptoms. *Journal of Affective Disorders*, *193*, 208–214. https://doi.org/10.1016/j.jad.2015.12.047
Wittchen, H. U., Jacobi, F., Rehm, J., Gustavsson, A., Svensson, M., Jönsson, B., … Steinhausen, H.-C. (2011). The size and burden of mental disorders and other disorders of the brain in Europe 2010. *European Neuropsychopharmacology*, *21*(9), 655–679. https://doi.org/10.1016/j.euroneuro.2011.07.018
Zald, D. H., & Treadway, M. T. (2017). Reward processing, neuroeconomics, and psychopathology. *Annual Review of Clinical Psychology*, *13*(1), 471–495. https://doi.org/10.1146/annurev-clinpsy-032816-044957

# Narratives, environments, and decision-making: A fascinating narrative, but one to be completed

Julian N. Marewski

Faculty of Business and Economics, Department of Organizational Behavior, University of Lausanne, 1015 Lausanne, Switzerland.
Julian.marewski@unil.ch
https://hecnet.unil.ch/hec/recherche/fiche?pnom=jmarewski&dyn_lang=en

doi:10.1017/S0140525X22002564, e102

**Abstract**

I encourage Johnson et al. to ground Conviction Narrative Theory in more detail in foundational, earlier decision-making research – first and foremost in Herbert Simon's work. Moreover, I wonder if and how further reflections about narratives could aid tackling two interrelated grand challenges of the decision sciences: To describe decision-making environments; to understand how people select among decision-strategies in environments.

Conviction Narrative Theory stresses how important narratives are for decision-making. I agree, and would encourage Johnson et al. to anchor Conviction Narrative Theory more in foundational work on decision-making – first and foremost in that of Herbert Simon.

Moreover, I believe reflecting about narratives could aid tackling two interrelated theoretical challenges of the decision sciences: How to describe decision-making environments; how to explain how people select among different decision-strategies for action as a function of their environment?

## *How to describe environments?*

Simon (1990) stressed that "Human rational behavior …is shaped by a scissors whose two blades are the structure of task environments and the computational capabilities of the actor" (p. 7). Together with Newell, he developed a research program that characterizes those blades.

Narratives could serve to capture and transmit actionable insights for tackling *ill-structured problems* (Simon & Newell, 1958). In such problems, "the objective function, the goal, is vague and nonquantitative" (p. 5), key variables are non-numerical, and "computational algorithms … are not available" (p. 5). Simon and Newell (1958) contrasted such problems to *well-structured* ones, which "can be formulated explicitly and quantitatively, and …be solved by known and feasible computational techniques" (p. 5). Their typology of well-structured versus ill-structured problems is orthogonal to the Knightian (Knight, 1921/1971) contrast between *uncertainty* and *risk* Johnson et al. focus on: Both ill-structured and well-structured problems can come with uncertainty. What matters for decision-making is not just uncertainty – and be it Donald Rumsfeld's surprising "*unknown unknowns*" (Mousavi & Gigerenzer, 2017, p. 363) – but also the degree and type of structure a problem exhibits. Most challenging real-life problems do not come with clearly delineated boundaries; rather, they are undefined, they can shape-shift and evolve dynamically over time, or be interchained with, and leading into, new problems. Particularly social task environments – which Johnson et al. discuss – exhibit such structural fuzziness.

## *How do people select among different strategies for action?*

One may speculate that narratives allow recognizing key elements of ill-structured problems. Recognizing one's task environment and patterns therein may help knowing what strategies one should rely on. Indeed, Simon (e.g., Simon, 1990; Simon & Chase, 1973) stressed how important recognition processes are for human decision-making and rationality. My hypothesis is that narratives serve what is known as *strategy selection* in the decision sciences (e.g., Marewski & Schooler, 2011).

Moreover, narratives may inform us not only about what we *should* do, but also about what we *can* do in the first place. The latter function ('the can') is a prerequisite for selection ('the should'): Narratives likely transmit actionable insights; they are vehicles for passing on heuristics and other decision-making strategies across generations; they aid filling our toolbox of strategies. Think of holy books, fairy tales, or folk wisdom: "Do not put all your eggs in the same basket" (Hafenbrädl, Waeger, Marewski, & Gigerenzer, 2016, p. 218), "Do to others as you would have them do to you" (*The Bible*, Luke, 6:31), and other well-known simple guiding principles are heuristic decision-strategies (e.g., Marewski & Hoffrage, 2020). Polya (1945/2014) – a founding father of the study of heuristics whose insights shaped Newell's and Simon's work on heuristics (Dick, 2015) – saw how ancient proverbs and heuristics of more modern days correspond.

Many such decision tools have fuzzy scopes and their task environments fuzzy boundaries. Such lack of definiteness could be functional: One may speculate that fuzziness allows for those narrated descriptions of tools and environments to attach to different kinds of ill-structured real-world situations a decision-maker may experience.

Indeed, communication, selection, and pattern detection all form part of Conviction Narrative Theory; there may be room for an integrative research program.

### *Developing Conviction Narrative Theory's narrative*

Such a program warrants precise definitions of concepts concerning the environmental and mental blades, and those theoretical elements must fit to each other, analogously to how two blades must fit to cut. Currently, Johnson et al. invoke uncertainty to characterize environments only broadly. Future developments of Conviction Narrative Theory could specify different types of environments further, with structure being one dimension. Likewise, one could zoom into specific models of heuristics and organize the toolbox of different types of heuristics as a function of the description of environments available. For instance, most *fast-and-frugal heuristics* (Gigerenzer, Todd, & the ABC Research Group, 1999) are algorithmic models that can be applied to well-defined problems that feature uncertainty. Classification and probabilistic inference tasks are examples. Heuristics that come as qualitative guiding principles – as in proverbs – may help to manage diverse ill-defined tasks, with more fuzzy boundaries. For example, qualitative guiding principles such as the one described in Luke (6:31) may serve leadership and cooperation alike, and be it in business or in warfare (see also Marewski & Katsikopoulos, 2022).

Currently, Conviction Narrative Theory invokes heuristics too nonspecifically, brushing across Tversky and Kahneman's (1974) error-prone shortcuts and heuristics as ecologically grounded models of bounded rationality (e.g., Gigerenzer & Goldstein, 1996). Only the latter reflect Simon's (1990) scissors (see Petracca, 2021, for a discussion). Also accounts of more complex tools need detail. Bayesianism is an example: Johnson et al. do not distinguish between Lindley's (1983) "Universal Bayes" (Gigerenzer & Marewski, 2015, p. 431), allegedly applying to all environments, and Savage's (1954/1972) domain-specific Bayesian decision theory. The latter sets apart "small" and "grand" worlds (e.g., p. 84), complementing Knight's contrast between risk and uncertainty (see Binmore, 2017), and taking into account "human possibility" (p. 16), not too inconsistent with Simon's notion of bounded rationality. Future work could deep-dive into such earlier theories of decision-making, examining if and how they would fit Conviction Narrative Theory.

All of the above might translate into two tasks: (1) Theory integration, and (2) setting apart, in detail, Conviction Narrative Theory from alternative lines of thought. Tackling both tasks may aid to better transmit the narrative Conviction Narrative Theory itself represents, allowing researchers to better understand and work with that narrative.

### *Conclusion*

To conclude, many narratives feature heroes of the past. Newell and Simon's heuristics (e.g., 1956) set the grounds for an entire field. Heuristics and Simon's (e.g., 1955, 1956) notion of bounded rationality form part of Conviction Narrative Theory, albeit with relatively little reference to Simon, and so do Savage's (1954/1972) "small worlds" (e.g., p. 84), but without much reference to Savage. Polya and Newell are heroes of the past, currently still lost. Conviction Narrative Theory is a fascinating narrative, but one that remains to be completed.

**Acknowledgements.** I thank Ulrich Hoffrage for very helpful comments.

**Financial support.** This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

**Competing interest.** None.

## References

Binmore, K. (2007). Rational decisions in large worlds. *Annales d'Économie et de Statistique*, *86*, 25–41.
Dick, S. (2015). Of models and machines: Implementing bounded rationality. *Isis*, *106*, 623–634.
Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, *103*, 650– 669.
Gigerenzer, G., & Marewski, J. N. (2015). Surrogate science. The idol of a universal method for scientific inference. *Journal of Management*, *41*, 421–440.
Gigerenzer, G., Todd, P. M., & The ABC Research Group. (1999). *Simple heuristics that make us smart*. Oxford University Press.
Hafenbrädl, S., Waeger, D., Marewski, J. N., & Gigerenzer, G. (2016). Applied decision making with fast-and-frugal heuristics. *Journal of Applied Research in Memory and Cognition*, *5*, 215–231.
Knight, F. H. (1921/1971). *Risk, uncertainty and profit*. University of Chicago Press. (Original work published in 1921 by Houghton Mifflin Company.).
Lindley, D. V. (1983). Theory and practice of Bayesian statistics. *Journal of the Royal Statistical Society. Series D* (*The Statistician*), *32*, 1–11.
Marewski, J. N., & Hoffrage, U. (2020). The winds of change: The Sioux, Silicon Valley, society and simple heuristics. In R. Viale (Ed.), *Routledge handbook of bounded rationality* (pp. 280–312). Routledge.
Marewski, J. N., & Katsikopoulos, K. V. (2022). *Simon's scissors: Meta-heuristics for discovering heuristics*. Unpublished manuscript.
Marewski, J. N., & Schooler, L. J. (2011). Cognitive niches: An ecological model of strategy selection. *Psychological Review*, *118*, 393–437.
Mousavi, S., & Gigerenzer, G. (2017). Heuristics are tools for uncertainty. *Homo Oeconomicus*, *34*, 361–379.
Newell, A., & Simon, H. A. (1956). *The logic theory machine. A complex information processing system*. Paper presented at the Symposium on Information Theory, Cambridge, Massachusetts. Retrieved from https://exhibits.stanford.edu/feigenbaum/catalog/ct530kb5673 on April 30th, 2021.
Petracca, E. (2021). On the origins and consequences of Simon's modular approach to bounded rationality in economics. *The European Journal of the History of Economic Thought*, *28*, 708–732.
Polya, G. (1945/2014). *How to solve it. A new aspect of mathematical method*. Princeton University Press. (Original work published in 1945 by Princeton University Press.).
Savage, L. J. (1954/1972). *The foundation of statistics*. Dover Publications. (Original work published in 1954 by John Wiley & Sons.).
Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, *69*, 99–118.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, *63*, 129–138.
Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology*, *41*, 1–19.
Simon, H. A., & Chase, W. G. (1973). Skill in chess. *American Scientist*, *61*, 394–403.
Simon, H. A., & Newell, A. (1958). Heuristic problem solving: The next advance in operations research. *Operations Research*, *6*, 1–10.
*The Bible*. Retrieved from https://biblehub.com/niv/luke/6.htm (on February 11, 2023).
Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124–1131.

# Is Conviction Narrative Theory a theory of everything or nothing?

Ben R. Newell[a] and Aba Szollosi[b]

[a]School of Psychology, UNSW Sydney, NSW 2052, Australia and [b]Department of Psychology, University of Edinburgh, Edinburgh EH8 9JZ, UK.
ben.newell@unsw.edu.au; http://www2.psy.unsw.edu.au/Users/BNewell/
aba.szollosi@gmail.com

doi:10.1017/S0140525X22002679, e103

**Abstract**

We connect Conviction Narrative Theory to an account that views people as intuitive scientists who can flexibly create, evaluate, and modify representations of decision problems. We argue that without understanding how the relevant complex narratives (or indeed any representation, simple to complex) are themselves constructed, we also cannot know when and why people would rely on them to make choices.

At the conclusion of their thought-provoking paper, Johnson et al. provide some vivid imagery in an effort to contextualise their contribution. They contrast a view of Conviction Narrative Theory (CNT) as comprising "too much theoretical meat" (sect. 10.1, para. 2) with a view that the theory is "skeletal," containing too little substance. This tension between the meat and the bones, or the details versus the framework, runs throughout the target article and is never fully resolved – and so exactly what CNT offers to scholars of decision-making remains unclear.

At the heart of this tension lies an unresolved contradiction regarding the scope of the theory. In some contexts, Johnson et al. designate CNT as merely a theory of decision-making under radical uncertainty and *not* a theory of the many crucial aspects of cognition – explanation, analogy, causation, emotion – that are often abstracted away in classical approaches to judgement and decision-making. In other contexts, narratives are being offered as the all-encompassing "currency of thought." Determining how broadly CNT can contribute to our understanding necessitates consideration of how a decision unfolds.

The first task for any decision maker, as Johnson et al. note, is to understand their current situation. How people achieve this understanding is an issue that appears in various guises across cognitive science. The "frame problem" in artificial intelligence research, the "correspondence problem" in judgement and decision-making and the "problem of induction" in philosophy all ask how agents are able to identify and act on relevant aspects of a situation. To adapt one of Johnson et al.'s examples: How does the widower determine the provenance of the noise that awoke him and decide what to do?

Johnson et al.'s claim appears to be that the first thing the widower does is to realise that he is under conditions of radical uncertainty in which probabilities are unknowable. How this initial step is achieved is not specified by the theory but it seems crucial if the purview is decision-making under radical uncertainty and not *all* decision-making, or indeed deliberative behaviour in general.

If narratives are only invoked when probabilities do not apply then the theory needs to explain how people know they are in such a situation. If the claim is broader and that narratives, or perhaps the internal monologues that accompany them, are used as a general guide to thinking and deciding – the "currency of thought" – then there is a deep concern that the theory loses coherence and becomes a set of generic statements about various tenuously related cognitive and emotional processes. There is clearly value in pointing out the relevance of other research areas to the endeavour of understanding behaviour that is often viewed through the narrow lens of expected utility theory (EUT), but to claim that this broadening of perspective constitutes a novel theory seems like overreach.

Returning to our widower, the idea that he explicitly runs a probability calculus over the possible outcomes of different actions is, naturally, absurd. He has most likely got to this stage in his life without ever hearing of EUT or its implications – just like the vast majority of people. As such, asking whether or how he knows that he is under conditions of radical uncertainty is also nonsensical. He has heard a sound and has to make a decision. Invoking EUT to explain how he does this might provide a way to think about and analyse the potential consequences of different actions but it is not a psychological explanation of his behaviour (and was never intended to be, e.g., by von Neumann & Morgernstern, 1947).

We agree with Johnson et al. that narratives provide another way to think about this process, but just as probability-based theories can be critiqued for not explaining "where the numbers come from," narrative approaches, like CNT, suffer from underspecification of how narratives are constructed (e.g., Klayman, 2001; Newell, Lagnado, & Shanks, 2022). By focusing only on a narrow set of all possible representations (complex narratives), CNT is unable to account for when and why people might use simpler strategies to deal with problems. Johnson et al. claim that narratives are higher-order representations that flexibly include lower-order representations, but the threshold for lower versus higher-order representations remains murky, as does the mechanism for flexible integration. This is problematic because as long as we do not understand how the relevant narratives are themselves constructed (or indeed any representation, simple to complex), we also cannot know when and why people would rely on them to make choices.

To answer the when and why questions, we would typically turn to diagnostic empirical evidence but here again the picture is rather unclear. Johnson et al. have a somewhat conflicting view stating both that lab-experimentation is "ill-suited" for testing the prevalence of narrative thinking, and that "finer-grained experimental evidence" is required to draw more definitive conclusions than those derived on the basis of qualitative interviews.

One potential solution to these problems is to start more modestly and focus on how people build up simple representations. Not all situations necessitate complex, emotionally charged simulations and evaluations, but they always necessitate the development of some kind of representation. Considering people to be "intuitive scientists" puts the focus on the process of how representations are generated and not on a particular kind of complex representation (Szollosi & Newell, 2020; Szollosi, Donkin, & Newell, 2022). From this perspective, understanding how people develop relatively simple representations – such as of integer numbers (Carey & Barner, 2019) or of the frequency of various events (Mason, Szollosi, & Newell, 2022) – can provide a useful starting point by enabling clearer ways to measure or manipulate the range of factors that Johnson et al. (rightly) claim play into any decision.

Explaining how such representations emerge does not require probabilities, or utilities, or the kinds of calculations that CNT wants to eschew, and we expect that focussing at this simpler level could eventually provide a window into how more complex representations (narratives) are built from primitive information. Without thinking carefully about these initial steps of flexible representation generation and selection there is a risk that the theoretical "tastiness" of CNT will be lost in a soup of conceptual complexity.

**Financial support.** Funding from the Australian Research Council (DP1901675) and from the EPSRC (EP/T033967/1) is gratefully acknowledged.

**Competing interest.** None.

## References

Carey, S., & Barner, D. (2019) Ontogenetic origins of human integer representations. *Trends in Cognitive Sciences*, *23*, 823–835.
Klayman, J. (2001). Ambivalence in (not about) naturalistic decision making. *Journal of Behavioral Decision Making*, *14*, 372–373.
Mason, A., Szollosi, A., & Newell, B. R. (2022). Learning the lie of the land: How people construct mental representations of distributions. Preprint. Available at https://psyarxiv.com/rdhv8/
Newell, B. R., Lagnado, D. A., & Shanks, D. R. (2022). *Straight choices: The psychology of decision making* (3rd ed.). Routledge.
Szollosi, A., Donkin, C., & Newell, B. R. (2022). Toward nonprobabilistic explanations of learning and decision-making. *Psychological Review*. Advance online publication. https://doi.org/10.1037/rev0000355
Szollosi, A., & Newell, B. R. (2020). People as intuitive scientists: Reconsidering statistical explanations of decision making. *Trends in Cognitive Sciences*, *24*(12), 1008–1018.
von Neumann, J., & Morgernstern, O. (1947). *Theory of games and economic behaviour* (2nd ed.). Princeton University Press.

# Narratives of crisis: From affective structures to adaptive functions

Petra Pelletier[a,b], Cécile McLaughlin[b] and Magali Boespflug[c]

[a]Laboratory of Social Psychology (EA 4471), University of Paris, Paris 92774, France; [b]Research Center for Semiotics, CeReS (EA 3648), Faculty of Letters and Human Sciences (FLSH), University of Limoges, Limoges 87000, France and [c]Research Center for Management (CEREGE), Department of Legal, Economic and Management Sciences, University of Poitiers, Poitiers 86073, France.
petra.pelletier@gmail.com; petra.pelletier@unilim.fr
cecile.mc-laughlin@unilim.fr
magali.boespflug@univ-poitiers.fr

doi:10.1017/S0140525X22002618, e104

**Abstract**

This commentary focuses on affective structures and the main adaptive functions of shared narratives to fill the gaps of the Conviction Narrative Theory. The transmission of narratives among individuals in highly uncertain situations is irrevocably tainted by affects and anchored in collective memory. Narratives have important evolutionary functions for human beings under threat and act as the social glue that creates and strengthens social bonds among individuals.

Since the dawn of time, the transmission of narratives among individuals tends to proliferate in times of wars; natural and man-made disasters, the COVID-19 crisis; and other dangerous and highly uncertain situations. These narratives that allow people to construct meanings of the crisis and cope with adversity are irrevocably tainted by affects and anchored in collective memory (Seeger & Sellnow, 2016). Because the Conviction Narrative Theory (CNT), proposed by Johnson et al., focuses primarily on textual objects, we aim to highlight the extent to which the structural and generative semiotics can constitute a relevant complementary contribution to the choice of narratives deployed by individuals in situations of radical uncertainty.

The singularity of semiotics as an interdisciplinary research field is to grasp constructed or coded elements that bear complementarity to social and cognitive sciences to seize the in-depth meanings of human communication (Greimas, 1966). The semiotic analysis of narratives implies identifying a structure, a dynamic that renders it possible to understand the narratives' architecture and deep mechanics. The narratives' architecture stems from the narratives deployed; their rhetorical and argumentative arrangements; the identified discursive structures, denotations, and connotations; and, finally, the systems of values and emotions that underlie these narratives. The dynamics of a narrative arise from a network of tensions rooted in a lacunar situation, an initial lack or trigger that acts as a revealer. In particular, during the COVID-19 crisis, a meticulous analysis of narratives revealed not only the collectively shared imaginaries that reflect humans' primordial primary fears of the plague disease but also the collective martial imaginary (McLaughlin, Pelletier, & Boespflug, 2022).

Indeed, structured semiotic analysis reveals the architectures of meaning at the narrative level and facilitates understanding individuals' emotions underlying a narrative, an action, or a practice. Nevertheless, CNT, which evokes the affective dimension as a fundamental constituent of narrative thinking, does not consider any tools for its analyses, whereas the semiotics of passions, founded by Greimas and Fontanille (1991), allows one to seize the experience of the subject, a subject governed by their feelings. The passion, which is a crucial and prototypical component of the Paris School of Semiotics (Broden, 1992), requires a thorough textual analysis of narratives governed by sensitive, somatic, and perceptual dimensions. The semiotic analysis of passion narratives demonstrates that this universe of modalities and configurations precedes a signification, is situated at a pre-cognitive level, and is animated by tensions that reveal various valences, areas of attraction and repulsion, and "states of mind." These antagonistic narrative configurations affirm the primary affective and social nature of human beings that tends to transcend historical and cultural boundaries (Greimas & Fontanille, 1991).

Narratives have other prominent functions for human beings in addition to those mentioned in the target article. First, from the evolutionary perspective, narratives represent a universal activity that allows the transmission of survival-relevant information and meaning-making of new, uncertain, and dangerous objects or situations (Pelletier & Drozda-Senkowska, 2020). Second, narratives are transmitted to provide audiences with relevant information that allows them to construct representations of an unknown object or situation. Thus, narratives might be particularly persuasive in manipulating the audiences' representations of the physical or social environment to transmit or create desired reality (Sugiyama, 1996, 2017). Third, direct, person-to-person transmission of narratives tends to create and strengthen social bonds among individuals and, therefore, facilitates social cohesion and cooperation, which are elemental for individuals' survival (Bietti, Tilston, & Bangerter, 2019).

Thus, although people may be unaware of the affective structures and adaptive functions of narratives, especially in the context of radical uncertainty, there seems to be no doubt about their functions in terms of creating reality and strengthening interpersonal relationships. This offers a solid reason for strategically selecting one's narratives, which equates to persuasively manipulating reality and intersubjectivity. Thus, narratives are a powerful social construction tool that depends on the manner and the social context in which a discourse is constructed by individuals to structure the shapeless uncertainty.

**Financial support.** This research was conducted as a part of the research program CoviZion "Storytelling of a Virus: Representations, Images and Imaginations. Represent and Understand in Order to Act Better and Live with it," funded by AMI FLASH Research and Innovations COVID of Nouvelle-Aquitaine, France.

**Competing interest.** None.

## References

Bietti, L. M., Tilston, O., & Bangerter, A. (2019). Storytelling as adaptive collective sensemaking. *Topics in Cognitive Science*, *11*(4), 710–732. https://doi.org/10.1111/tops.12358

Broden, T. F. (1992). Paris school passions: Theory and practice. *Semiotics*, 27–33. https://doi.org/10.5840/cpsem19927

Greimas, A. J. (1966). *Sémantique structurale: Recherche de méthode. [Structural Semantics: An Attempt at a Method]*. Larousse.

Greimas, A. J., & Fontanille, J. (1991). *Sémiotique des passions: Des états de choses aux états d'âme. [Semiotics of passions: From states of affairs to states of feelings]*. Seuil.

McLaughlin, C., Pelletier, P., & Boespflug, M. (2022). Storytelling of a virus: A focus on COVID-19 narratives of older adults. *International Journal of Arts, Humanities and Social Studies*, *4*(1), 86–95.

Pelletier, P., & Drozda-Senkowska, E. (2020). Towards a socially situated rumouring: Historical and critical perspectives of rumour transmission. *Social and Personality Psychology Compass*, *14*(6), e12532. https://doi.org/10.1111/spc3.12532

Seeger, M. W., & Sellnow, T. L. (2016). *Narratives of crisis: Telling stories of ruin and renewal*. Stanford University Press.

Sugiyama, M. S. (1996). On the origins of narrative: Storyteller bias as a fitness-enhancing strategy. *Human Nature*, *7*(4), 403–425. https://doi.org/10.1007/BF02732901

Sugiyama, M. S. (2017). Oral storytelling as evidence of pedagogy in forager societies. *Frontiers in Psychology*, *8*, 471. https://doi.org/10.3389/fpsyg.2017.00471

# What makes narratives feel right? The role of metacognitive experiences

Norbert Schwarz

Department of Psychology, University of Southern California, Los Angeles, CA 90089-1061, USA.

Norbert.Schwarz@usc.edu
https://dornsife.usc.edu/norbert-schwarz/

doi:10.1017/S0140525X22002758, e105

**Abstract**

Conviction Narrative Theory holds that reasoners adopt "a narrative that feels 'right' to explain the available data" and use "that narrative to imagine plausible futures" (target article, Abstract). Drawing on feelings-as-information theory, this commentary reviews the role of metacognitive experiences of ease or difficulty and highlights that fluently processed narratives are more likely to "feel right."

Central to Johnson et al.'s Conviction Narrative Theory (CNT) is the assumption that reasoners adopt "a narrative that feels 'right' to explain the available data" and use "that narrative to imagine plausible futures" (target article, Abstract). Their discussion of what makes a narrative "feel right" focuses on structural aspects of the narrative and neglects the role of metacognitive experiences in the construction and evaluation of narratives. This commentary addresses this gap, drawing on feelings-as-information theory (Schwarz, 2012; Schwarz & Clore, 2007).

Narratives are more compelling when they are easy to construct. People monitor the dynamics of their own thought processes and have more confidence in their thoughts when processing is easy (Koriat, 2007; Schwarz, 2015). Numerous content, context, and person variables can influence the subjective experience of fluent processing (Reber, Schwarz, & Winkielman, 2004; Schwarz, Jalbert, Noah, & Zhang, 2021). Some influences are integral to what the person thinks about (e.g., the coherence or complexity of a narrative), whereas others are incidental (e.g., momentary distractions or the readability of a print font). Because people are more sensitive to their feelings than to the source of their feelings, they frequently misread incidental feelings as bearing on what they are thinking about. The inferences they draw from processing fluency are guided by lay theories of mental operations, which are usually correct. People assume, for example, that a narrative is easier to process when it is familiar, coherent, and compatible with other things they know than when it is not. Hence, they infer higher familiarity (Song & Schwarz, 2009) and coherence (Topolinski & Strack, 2009) when processing is easy, and are less likely to notice a claim's incompatibility with their own knowledge (Song & Schwarz, 2008), even when the experience is solely due to variations in color contrast, print font, or ease of pronunciation. Because familiarity is an important input into judgments of risk and trust, people are also more likely to invest in stocks when the ticker symbol is easy to pronounce (Alter & Oppenheimer, 2006; Green & Jame, 2013), indicating reduced risk perception, and to trust online sellers when their usernames are easy to process (Silva, Chrobot, Newman, Schwarz, & Topolinski, 2017). Moreover, fluent processing due to repetition (Hasher, Goldstein, & Toppino, 1977), cultural familiarity (Oyserman, 2019), audio quality (Bild et al., 2021), color contrast (Reber & Schwarz, 1999), rhyme (McGlone & Tofighbakhsh, 2000), related nonprobative photos (Newman & Zhang, 2021), and many other incidental variables increases the acceptance of claims as true (Brashier & Marsh, 2020; Schwarz & Jalbert, 2021).

This suggests that only fluent narratives are compelling enough to elicit strong "convictions," whereas disfluent narratives are likely to seem implausible (Schwarz, Sanna, Skurnik, & Yoon, 2007). Debiasing research supports this conclusion. After an outcome is known, people believe that they could have predicted it (Fischhoff, 1975). To attenuate this hindsight bias, it is often recommended to consider how the event might have turned out otherwise (Larrick, 2004), thus prompting the development of an alternative narrative. Because thinking of just a few reasons for alternative outcomes is more difficult than thinking of many, this debiasing strategy is successful when people are asked for a few reasons, but backfires when they are asked for many (Sanna, Schwarz, & Stocker, 2002). Even when people successfully generate many reasons for alternative outcomes, the associated difficulty convinces them that the obtained outcome was almost unavoidable. All three results – the emergence of a hindsight bias, its attenuation through generating a few alternatives, and its exaggeration by generating many alternatives – reflect judges' reliance on their own metacognitive experiences and can be obtained with different incidental and integral manipulations of processing fluency (Schwarz et al., 2007).

In addition to providing information about one's processing dynamics, fluent processing is experienced as pleasant and elicits a spontaneous positive affective response that can be captured with psychophysiological measures (Winkielman & Cacioppo, 2001). Hence, a given stimulus is liked more when it is easy to process, for example, due to repetition (Zajonc, 1968), visual (Reber, Winkielman, & Schwarz, 1998) or conceptual primes (Belke, Leder, Strobach, & Carbon, 2010), or individual differences in color vision (Álvaro, Moreira, Lillo, & Franklin, 2015). This affective pathway from fluent processing to liking parallels the influence of happy and sad moods (Schwarz, 1990) and figures prominently in aesthetic preference and judgments of beauty (Reber et al., 2004). The common intuition that beauty and truth go hand in hand reflects that processing fluency serves as an input into both judgments (Schwarz, 2018).

CNT holds that "explanatory fit is experienced affectively" (sect. 6.2). Processing fluency provides affective as well as non-affective information that informs assessments of what one is thinking about – when thoughts flow smoothly, people nod along (Schwarz & Jalbert, 2021). The observed parallel influence of integral and incidental variables implies that assessments of explanatory fit are highly error prone – because people are more sensitive to their feelings than to the source of their feelings, narratives can feel "right" (or "wrong") for reasons that are substantively unrelated to the content of the narrative.

For researchers, this source insensitivity provides two useful methodological tools. In Johnson et al.'s examples, narrative content and associated feelings cannot be separated – simulating a narrative that leads to a positive (negative) future elicits positive (negative) feelings and either the content of the narrative or the feeling it elicits may drive subsequent choices. To separate the contributions of content and feelings, one can (1) elicit an incidental feeling, as in the above examples, or (2) undermine the informational value of integral feelings by leading people to attribute their feelings to an incidental source, which attenuates or eliminates the impact of feelings (e.g., Novemsky, Dhar, Schwarz, & Simonson, 2007; Schwarz & Clore, 1983). Without such manipulations, CNT's causal claims about the roles of narrative content versus the accompanying (affective or non-affective) feelings, and their relative contributions at different stages, remain ambiguous. The further development of CNT will benefit from tackling this task.

**Acknowledgement.** None.

**Financial support.** No funding for this commentary.

**Competing interest.** None.

## References

Alter, A. L., & Oppenheimer, D. M. (2006). Predicting short-term stock fluctuations by using processing fluency. *Proceedings of the National Academy of Sciences*, *103*(24), 9369–9372.

Álvaro, L., Moreira, H., Lillo, J., & Franklin, A. (2015). Color preference in red–green dichromats. *Proceedings of the National Academy of Sciences*, *112*, 9316–9321.

Belke, B., Leder, H., Strobach, T., & Carbon, C. C. (2010). Cognitive fluency: High-level processing dynamics in art appreciation. *Psychology of Aesthetics, Creativity, and the Arts*, *4*(4), 214–222.

Bild, E., Redman, A., Newman, E. J., Muir, B., Tait, D., & Schwarz, N. (2021). Sound and credibility in the virtual court: Low audio quality leads to less favorable evaluations of witnesses and lower weighting of evidence. *Law & Human Behavior*, *45*(5), 481–495.

Brashier, N. M., & Marsh, E. J. (2020). Judging truth. *Annual Review of Psychology*, *71*(1), 499–515.

Fischhoff, B. (1975). Hindsight 6¼ foresight: The eVect of outcome knowledge on judgments under uncertainty. *Journal of Experimental Psychology: Human Perception and Performance*, *1*, 288–299.

Green, T. C., & Jame, R. (2013). Company name fluency, investor recognition, and firm value. *Journal of Financial Economics*, *109*(3), 813–834.

Hasher, L., Goldstein, D., & Toppino, T. (1977). Frequency and the conference of referential validity. *Journal of Verbal Learning and Verbal Behavior*, *16*, 107–112.

Koriat, A. (2007). Metacognition and consciousness. In P. D. Zelazo, M. Moscovitch & E. Thompson (Eds.), *The Cambridge handbook of consciousness* (pp. 289–326). Cambridge University Press.

Larrick, R. P. (2004). Debiasing. In D. J. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 316–337). Blackwell Publishing.

McGlone, M. S., & Tofighbakhsh, J. (2000). Birds of a feather flock conjointly (?): Rhyme as reason in aphorisms. *Psychological Science*, *11*(5), 424–428.

Newman, E. J., & Zhang, L. (2021). Truthiness: How non-probative photos shape belief. In R. Greifeneder, M. Jaffé, E. J. Newman & N. Schwarz (Eds.), *The psychology of fake news: Accepting, sharing, and correcting misinformation* (pp. 90–114). Routledge/ Psychology Press.

Novemsky, N., Dhar, R., Schwarz, N., & Simonson, I. (2007). Preference fluency in choice. *Journal of Marketing Research*, *44*, 347–356.

Oyserman, D. (2019). Cultural fluency, mindlessness, and gullibility. In R. Baumeister & J. P. Forgas (Eds.), *The social psychology of gullibility: Conspiracy theories, fake news and irrational beliefs* (pp. 255–278). Routledge.

Reber, R., & Schwarz, N. (1999). Effects of perceptual fluency on judgments of truth. *Consciousness and Cognition*, *8*(3), 338–342.

Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review*, *8*, 364–382.

Reber, R., Winkielman, P., & Schwarz, N. (1998). Effects of perceptual fluency on affective judgments. *Psychological Science*, *9*(1), 45–48.

Sanna, L. J., Schwarz, N., & Stocker, S. L. (2002). When debiasing backfires: Accessible content and accessibility experiences in debiasing hindsight. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 497–502.

Schwarz, N. (1990). Feelings as information: Informational and motivational functions of affective states. In E. T. Higgins & R. M. Sorrentino (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (Vol. *2*, pp. 527–561). Guilford Press.

Schwarz, N. (2012). Feelings-as-information theory. In P. A. M. Van Lange, A. Kruglanski & E. T. Higgins (Eds.), *Handbook of theories of social psychology* (pp. 289–308). Sage.

Schwarz, N. (2015). Metacognition. In M. Mikulincer, P. R. Shaver, E. Borgida & J. A. Bargh (Eds.), *APA handbook of personality and social psychology: Attitudes and social cognition* (pp. 203–229). APA.

Schwarz, N. (2018). Of fluency, beauty, and truth: Inferences from metacognitive experiences. In J. Proust & M. Fortier (Eds.), *Metacognitive diversity. An interdisciplinary approach* (pp. 25–46). Oxford University Press.

Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology*, *45*, 513–523.

Schwarz, N., & Clore, G. L. (2007). Feelings and phenomenal experiences. In A. Kruglanski & E. T. Higgins (Eds.), *Social psychology: Handbook of basic principles* (2nd ed., pp. 385–407). Guilford Press.

Schwarz, N., & Jalbert, M. C. (2021). When fake news feels true. Intuitions of truth and the acceptance and correction of misinformation. In R. Greifeneder, M. Jaffé, E. J. Newman & N. Schwarz (Eds.), *The psychology of fake news: Accepting, sharing, and correcting misinformation* (pp. 73–90). Routledge/Psychology Press.

Schwarz, N., Jalbert, M. C., Noah, T., & Zhang, L. (2021). Metacognitive experiences as information: Fluency in consumer judgment and decision making. *Consumer Psychology Review*, *4*(1), 4–25.
Schwarz, N., Sanna, L., Skurnik, I., & Yoon, C. (2007). Metacognitive experiences and the intricacies of setting people straight: Implications for debiasing and public information campaigns. *Advances in Experimental Social Psychology*, *39*, 127–161.
Silva, R. R., Chrobot, N., Newman, E., Schwarz, N., & Topolinski, S. (2017). Make it short and easy: Username complexity determines trustworthiness above and beyond objective reputation. *Frontiers in Psychology*, *8*, 2200.
Song, H., & Schwarz, N. (2008). Fluency and the detection of misleading questions: Low processing fluency attenuates the Moses illusion. *Social Cognition*, *26*(6), 791–799.
Song, H., & Schwarz, N. (2009). If it's difficult to pronounce, it must be risky: Fluency, familiarity, and risk perception. *Psychological Science*, *20*, 135–138.
Topolinski, S., & Strack, F. (2009). The architecture of intuition: Fluency and affect determine intuitive judgments of semantic and visual coherence and judgments of grammaticality in artificial grammar learning. *Journal of Experimental Psychology: General*, *138*, 39–63.
Winkielman, P., & Cacioppo, J. T. (2001). Mind at ease puts a smile on the face: Psychophysiological evidence that processing facilitation leads to positive affect. *Journal of Personality and Social Psychology*, *81*, 989–1000.
Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, *9*(2, Pt. 2), 1–27.

# Educational Implications of Conviction Narrative Theory

Mark Sheskin, Michael Bogucki, Tomer Perry and Katie McAllister

Minerva University, San Francisco, CA 94103, USA.
msheskin@minerva.edu
mbogucki@minerva.edu
tperry@minerva.edu
kmcallister@minerva.edu
https://www.marksheskin.com/

doi:10.1017/S0140525X22002631, e106

**Abstract**

Education often relies on an implicit assumption that decisions are made rationally, and focuses on situations in which there are correct answers that can be known with certainty. The proposal that decision-making is often narrative, especially in contexts of radical uncertainty, suggests important changes to education practice and new questions for education research.

We outline three educational implications of Conviction Narrative Theory (CNT).

First, CNT can contribute to understanding pre-existing results in the education literature, especially regarding challenges to effective learning that is retained and applied in new contexts (e.g., Barnett & Ceci, 2002). As a concrete example, imagine a biology teacher who hopes that class content will help students understand the world and make better decisions in the future. Perhaps, years after the class, students might be able to correctly identify a new conspiracy theory around vaccines as being biologically implausible. Such an achievement requires that relevant information is initially understood and encoded into memory, remembered over a long time, recalled unprompted and outside of the original learning context, and finally applied correctly to a novel context ("far transfer"). CNT makes clear some of the challenge: When people are constructing potential narratives, snippets of information learned for a very narrow goal and applied in a very narrow context long ago (e.g., memorizing what "Helper T cells" are, to pass a biology exam about the immune system) may have very low salience as building blocks for narratives.

Second, CNT has implications for educational practice: Teachers can help students identify when intuitive narrative is *or is not* best suited to understanding and decision-making. When narrative is useful, teachers can help students learn to use it more effectively.

Consider an example of narratives interfering with reasoning, such that increased knowledge paradoxically increases wrong judgments. Kahan et al. (2012) found that "Members of the public with the highest degrees of science literacy and technical reasoning capacity… were the ones among whom cultural polarization was greatest" (e.g., about climate change). Such results indicate the power of the "narratives": Teaching and learning more "facts" can be useless – or even harmful – if incorporated into a flawed narrative framework. This is consistent with research on confirmation bias (Nickerson, 1998), but CNT makes clear *what* is being confirmed in many of these cases: A particular narrative about the world. Regarding education, it suggests that students might be guided to be particularly cautious about narratives when they feel strongly about a topic but don't yet know a lot about it.

In contrast, other examples show narratives being an asset for learning and reasoning. Dudley and Bold (1996) revised a non-calculus-based introductory physics course to start with everyday life (e.g., how does electricity within a house work?) and then worked down to the low-level physics phenomena (e.g., voltage). Notably, this approach "does not insist on mathematical rigor at every step, but concentrates more on the development of conceptual understanding."

These contrasting examples illustrate the difficulty of knowing when to employ narrative. Indeed, some colleagues objected to the revised physics course because "they fear that any neglect of rigor compromises the physics education of the students." This gap between what instructors think is best and what works well for students may be partially explained by "the curse of knowledge": Once you know something, it is hard to simulate the mind of someone who does not know it (Wieman, 2007). Imagine a psychology teacher explaining the proper interpretation of $p$-values. Even when the lesson focuses on the mathematical level, the teacher's rich understanding of $p$-values is embedded in knowledge of the "replication crisis" in psychology. She might be thinking about Simmons, Nelson, and Simonsohn (2016) reporting that "undisclosed flexibility in data collection and analysis allows presenting anything as significant"; or Bennett, Miller, and Wolford (2009) finding "results" from an fMRI of a dead salmon; or check-lists to avoid $p$-hacking (Wicherts et al., 2016); or "meta-discussion" about the tone around discussions of replication (Derksen & Field, 2022). In short, the instructor has a huge amount of related knowledge available to contextualize – and narrativize – her understanding of the low-level details being discussed; the students do not, and this gap can make pedagogy extremely challenging.

Third, CNT suggests priorities for future education research. Future research can identify both how to best use narratives *and* how to best teach students to use narratives. Consider simulation games that place students in the heart of the action. Shared experience of students in role-playing exercises becomes a useful narrative that the instructor can leverage to translate knowledge to terms meaningful to students. For example, one of us has had students do a classroom exercise simulating the working of direct democracy, with students taking roles with different interests and knowledge in a self-governing democratic city facing a momentous decision. By pursuing their interests in the exercise, students end up experiencing a variety of important social and political phenomena, such as the spread of misinformation. Class comes "built-in" with the narrative of what happened in the exercise that led to the spread of misinformation, providing the instructor with a concrete narrative for analysis to ensure a shared understanding of the concepts and theories. Such practices seem well-supported by various aspects of CNT, but future research can help establish which practices actually lead to the best learning outcomes.

As an additional consideration, how can teachers give clear feedback when the focus is not on whether the final answer is "correct" but rather the process (and narrative) leading to the answer? Teachers must challenge students to break apart and critically examine the narratives underlying decisions and perspectives. This approach presents an opportunity to teach a transferable skill rather than discipline-specific knowledge. Through an evaluative lens: Are the students presenting structured evidence to justify their position? Is this evidence of high source quality? Do valid causal links tie narrative elements? Have they identified and mitigated biases that could have undue influence? Such questions prompt faculty to craft feedback aligned with Hattie and Timperley's task, process, and self-regulatory levels (2007) that is likely to serve as "high-information feedback," shaping learner behavior in more productive ways than simple reinforcement or punishment (Wisniewski, Zierer, & Hattie, 2020). This deeper engagement with process can enable students to regularly sense-check their biases and assumptions, be more open to incorporating new evidence and updating perspectives, and be more resistant to superficially compelling yet unfounded narratives.

**Financial support.** This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

**Competing interest.** None.

## References

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn?: A taxonomy for far transfer. *Psychological Bulletin*, *128*(4), 612.
Bennett, C. M., Miller, M. B., & Wolford, G. L. (2009). Neural correlates of interspecies perspective taking in the post-mortem Atlantic Salmon: An argument for multiple comparisons correction. *Neuroimage*, *47*(Suppl 1), S125.
Derksen, M., & Field, S. (2022). The tone debate: Knowledge, self, and social order. *Review of General Psychology*, *26*(2), 172–183.
Dudley, J. M., & Bold, G. E. (1996). "Top-down teaching" in noncalculus-based introductory physics classes. *American Journal of Physics*, *64*(4), 418–421.
Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*(1), 81–112.
Kahan, D. M., Peters, E., Wittlin, M., Slovic, P., Ouellette, L. L., Braman, D., & Mandel, G. (2012). The polarizing impact of science literacy and numeracy on perceived climate change risks. *Nature Climate Change*, *2*(10), 732–735.
Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*(2), 175–220.
Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2016). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science, 22*(11), 1359–1366.
Wicherts, J. M., Veldkamp, C. L., Augusteijn, H. E., Bakker, M., Van Aert, R., & Van Assen, M. A. (2016). Degrees of freedom in planning, running, analyzing, and reporting psychological studies: A checklist to avoid p-hacking. *Frontiers in Psychology*, *7*, 1832.
Wieman, C. E. (2007). APS News–The back page. The "curse of knowledge" or why intuition about teaching often fails. *American Physical Society News*, *16*(10).
Wisniewski, B., Zierer, K., & Hattie, J. (2020). The power of feedback revisited: A meta-analysis of educational feedback research. *Frontiers in Psychology*, *10*, 3087.

# Making the unconscious conscious: Developing maladaptive scripts into conviction narratives

Paul Siegel

Department of Psychology, Purchase College, State University of New York, Purchase, NY 10577, USA.
paul.siegel@purchase.edu
https://www.purchase.edu/live/profiles/671-paul-siegel

doi:10.1017/S0140525X22002734, e107

**Abstract**

In his 'script theory,' Tomkins first proposed that people unconsciously organize their life experiences in terms of narrative structures he termed "scripts." I use a clinical vignette to illustrate how the psychotherapeutic process of "making the unconscious conscious" involves becoming aware of the maladaptive scripts that people unwittingly live by, and developing them into the "conviction narratives" proposed by the authors.

A psychotherapy patient arrived late, hurried into my office, and threw herself down in her seat. Tears had ruined her eye makeup and created black streams down her cheeks, although her facial expression was furious. As our eyes met, she declared with a force that slapped me across the face, "I hate drunks." The patient was referring to her husband, whose alcohol dependence made life painful for her and their only child. A few seconds after I absorbed the impact of her fury and asked what had happened, nascent thoughts coalesced into a singular question in my mind: "If you hate drunks so much, why did you marry one?"

In fact, a familiar story was played out in her decision to marry this man – a kind, high-functioning survivor of childhood sexual abuse who, like the patient's father, drowned his emotional pain in alcohol. She genuinely loved her husband and wanted to heal him, a wish that as a child she had been able to fulfill, if only intermittently, in relation to her father. Yet she lacked emotional insight into the connection between her childhood relationship with her father and her marital choice. An unconscious script ran through her relationships with these two men, which can be represented as:

> He is in pain; I am forlorn → I try to ease his pain → He feels better and engages with me → I feel connected to him

This structure illustrates a more automatic and unconscious level of narrative processing than the "conviction narratives" proposed by the authors. The seminal personality theorist Silvan Tomkins first proposed that people employ narrative processing to organize their life experiences with his "script" theory (1979, 1987, 1991), which augured the rise of the narrative Zeitgeist across many fields of psychology (Demorest, Popovska, & Dabova, 2012; Howard, 1991; McAdams, 2001; Nelson & Fivush, 2004). To understand and cope with emotionally intense experiences, Tomkins proposed that people form "scripts" or affective schemas that represent how such experiences typically proceed and how to respond in order to maximize positive affect and minimize negative affect. While scripts are constructed from specific emotional experiences, they become general rather than specific so as to function as a personal guide for anticipating and managing similar experiences in the future. As Demorest et al. (2012) described, "The individual is both dramatist and actor, constructing the scripts through which he or she understands and lives life."

In this sense of providing an emotional guide, scripts are similar to the authors' conviction narratives. However, the latter are more complex, higher-order representations that allow people to imagine multiple possible futures – not just one, and evaluate those futures in arriving at consciously considered choices and action plans. Moreover, scripts are overlearned through repeated reinforcement and thus operate automatically and largely unconsciously, making them resistant to change. While originally adaptive, scripts can therefore become maladaptive when subsequently activated by ostensibly similar experiences that they do not, in fact, apply to. As a child, my patient's script allowed her to negotiate her father's pain and alcoholism and sustain a close relationship with him. Similarly becoming her husband's emotional caretaker, however, left her feeling exploited, exhausted, and entirely unmet in her marriage.

If a goal of psychotherapy is "to make the unconscious conscious," people need to become aware of the maladaptive scripts that they unwittingly live by, and develop them into reality-based conviction narratives. When my patient became aware of her script, she recognized that she *had* to take care of her father as a little girl in order to maintain a relationship with him. Similarly attending to her husband, and obviating her own needs in the process, however, need no longer be the price of intimacy. Indeed, continuing down this path would only destroy her marriage. So she decided on a different course. She would talk with her husband about how his drinking was impacting on their marriage and their son, and insist that he take responsibility for his illness by getting treatment and specialized social support. Only then could they relate to each other as marital partners and as parents, rather than as caretaker and patient. By developing fully considered narratives that accorded with the distinct emotional realities of her relationships with her father and husband, my patient was now free to imagine alternative futures and choose her own story, rather than repeat the painful story of her childhood.

The client's permission to write about her psychotherapy was obtained on tape more than 25 years ago, using general descriptors to make her and others unidentifiable.

**Financial support.** This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

**Competing interest.** None.

## References

Demorest, A., Popovska, A., & Dabova, M. (2012). The role of scripts in personal consistency and individual differences. *Journal of Personality*, *80*(1), 187–218.

Howard, G. S. (1991). Culture tales: A narrative approach to thinking, cross-cultural psychology, and psychotherapy. *American Psychologist*, *46*, 187–197.

McAdams, D. (2001). The psychology of life stories. *Review of General Psychology*, *5*, 100–122.

Nelson, K., & Fivush, R. (2004). The emergence of autobiographical memory: A social cultural developmental theory. *Psychological Review*, *111*, 486–511.

Tomkins, S. (1979). Script theory: Differential magnification of affects. In C. Keasey (Ed.), *Nebraska symposium on motivation* (Vol. *26*, pp. 201–236). University of Nebraska Press.

Tomkins, S. (1991). *Affect, imagery, consciousness* (Vol. *3*). Springer.

Tomkins, S. S. (1987). Script theory. In J. Aronoff, A. Rabin, & R. Zucker (Eds.), *The emergence of personality* (pp. 147–216). Springer.

# The statistical mechanics of felt uncertainty under active inference

Mark Solms

Department of Psychology, University of Cape Town, Rondebosch 7701, South Africa.
mark.solms@uct.ac.za

doi:10.1017/S0140525X22002783, e108

### Abstract

Convincing narratives are not confabulations. Presumably they "feel right" to decision-making agents because the probabilities they assign *intuitively* (i.e., implicitly) to potential outcomes are plausible. Can we render explicit the calculations that would be performed by a decision-making agent to evaluate the plausibility of competing narratives? And if we can, what, exactly, makes a narrative "feel right" to an agent?

Conviction Narrative Theory (CNT) claims that probabilistic formalisms cannot explain decision-making under radical uncertainty, where "probabilities cannot be assigned to outcomes." Accordingly, it replaces statistical analysis with a folk-psychological account: Decision-making agents adopt narratives that "feel right" to explain the available data and to imagine and evaluate plausible futures.

Since one narrative is more convincing than another, they presumably both encode probable outcomes *intuitively* – which is why the narrative that feels *right* is adopted. If so, what, exactly, makes a candidate narrative "feel right" to an agent?

I have argued elsewhere (Solms, 2021; Solms & Friston, 2018) that feelings represent fluctuating uncertainty under active inference: Decreasing uncertainty (inverse precision) in a course of action feels "good" for an agent because it increases confidence in its likelihood to minimize its variational free energy, and vice versa. It is possible to reduce this decision-making process to an equation: One that has recently been translated into folk-psychological terms by Smith, Ramstead, and Kiefer (2022). In what follows, therefore, I paraphrase them closely.

The relevant formalism (the POMDP formulation of active inference) pivots on the concept of a "policy" (denoted by $\pi$). A policy is a *possible course of action that decision-making agents can entertain*; it predicts a specific sequence of state transitions. What CNT calls "adopting a *narrative*," therefore, is formalized under active inference as "*policy* selection." Another pivotal concept is "expected confidence" (denoted by $\gamma$) which is a precision estimate for beliefs about expected free energy. "Expected free energy" (denoted by $G$) is the *future* variational free energy (the future $F$) expected under the outcomes anticipated when following a policy, given a model.

Here is the formalism:

$$G(\pi) = D_{KL}[q(o_\tau|\pi)||p(o_\tau)] + Eq(s_\tau|\pi)[H[p(o_\tau|s_\tau)]]$$

It says that a policy is more likely to be selected if it minimizes *risk* plus *ambiguity*.

Technically, under active inference, "risk" encodes the anticipated Kullback–Leibler divergence (denoted by $D_{KL}$) between two quantities. The first quantity ($q(o_\tau|\pi)$) corresponds to the observations an agent anticipates at each time-point (denoted by $o_\tau$) if it selects one policy versus another. This formalizes the folk-psychological notion of *expectations*. The second quantity ($p(o_\tau)$) corresponds to a policy-independent prior over observations which encodes the observations that are congruent with an agent's "prior preference distribution." This term formalizes the folk-psychological notion of *goals*. Selecting a policy to minimize free energy ($G(\pi)$) entails minimizing the difference between these prior expectations and the anticipated outcomes. An agent tries to infer which policy will generate outcomes closest to its goals, and selects the one that it *believes* is most likely to achieve what it *desires* (or *values*).

"Ambiguity," technically, encodes the expected entropy (denoted by $H$) of the likelihood function for a given state at each time-point (denoted by $s_\tau$). Entropy measures the precision of a distribution. This means that an agent that minimizes $G$ will actively seek states with the most precise mapping to observations (i.e., it will select policies expected to generate the most informative outcomes).

The posterior probability distribution over policies is

$$p(\pi) = \sigma(\ln E(\pi) - F(\pi) - \gamma G(\pi))$$

This says that the most likely policies are those which minimize $G$, under constraints afforded by $E(\pi)$ and $F(\pi)$ – where $\sigma$ denotes a function that converts the result back to a proper distribution with non-negative values summing to one. $E(\pi)$ encodes a fixed prior over policies that can be used to model *habitual* actions, in line with past influences. $F(\pi)$ scores the free energy of past and present observations under each policy (i.e., the evidence they provided for each policy). It reflects how well each policy predicts the observations that have already been received.

How much weight should an agent afford a policy based on *evidence from the past* relative to the expected free energy of *observations in the future*? It depends upon the precision estimate for beliefs about expected free energy ($\gamma$). This pivotal quantity controls how much model-based predictions about outcomes contribute to policy selection. It formalizes how much the predictions are *trusted*. Lower values for $\gamma$ cause an agent to act habitually and with less confidence in its future plans.

The precision estimate is updated with each new observation, allowing an agent to modulate confidence in its model of the future. This updating is performed via a hyperprior ($\beta$), the rate parameter of the $\gamma$ distribution:

$$p(\gamma) = \Gamma(1, \beta)$$

$$E[\gamma] = \gamma = 1/\beta$$

$$\beta_{update} \leftarrow \beta - \beta_0 + (p(\pi) - p(\pi_0)) \cdot (-G(\pi))$$

$$\beta \leftarrow \beta - \beta_{update}$$

$$\gamma \leftarrow 1/\beta$$

Here, the arrow ($\leftarrow$) indicates iterative value-updating (until convergence), $\beta_0$ is a prior on $\beta$, and $p(\pi_0)$ corresponds to $p(\pi)$ before an observation has been made to generate $F(\pi)$. That is, $p(\pi_0) = \sigma(\ln E(\pi) - \gamma G(\pi))$. For our purposes, this quantity within the value that updates the hyperprior ($\beta_{update}$) is especially important since it concerns *feelings*. This is a type of prediction error indicating whether a new observation provides evidence for or against beliefs about $G(\pi)$ – that is, whether or not $G(\pi)$ is consistent with the $F(\pi)$ generated by a new observation. When it leads to an increase in $\gamma$ (when confidence in $G(\pi)$ increases), it acts as evidence for a *positive* feeling, and vice versa.

The equation for posteriors over policies determines how likely a policy is. In folk-psychological terms, it determines an agent's *drive* to select one narrative over another. We have seen that this arises from two influences: The prior over policies ($E(\pi)$) and the expected free energy ($G(\pi)$). While $E(\pi)$ maps well to habitual policies, $G(\pi)$ reflects the inferred value of each policy based on beliefs (e.g., $p(o_\tau|\pi)$ and $p(o_\tau|s_\tau)$) and desired outcomes (i.e., $p(o_\tau)$). The policy with the lowest expected free energy therefore formalizes the *intentions* of an agent. In the absence of habitual influences, the intention would become the policy the agent *feels most driven to choose* in $p(\pi)$.

The pivotal role of feelings is clear. The formalism shows that when new observations support a current policy (when they are consistent with $G(\pi)$), $\gamma$ increases, which generates "good" feelings, and vice versa. This increase in $\gamma$ boosts an agent's confidence in its beliefs about expected variational free energy, which in turn reduces the appeal of habitual policies, $E(\pi)$.

In short, decision-making agents adopt *policies* that feel right to explain the available data and to imagine and evaluate plausible futures.

**Acknowledgement.** Ryan Smith kindly edited the MS for technical accuracy.

**Financial support.** This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

**Competing interest.** None.

## References

Smith, R., Ramstead, M., & Kiefer, A. (2022). Active inference models do not contradict folk psychology. *Synthese*, *200*, 81, https://doi.org/10.1007/s11229-022-03480-w

Solms, M. (2021). *The hidden spring: A journey to the source of consciousness*. W.W. Norton.

Solms, M., & Friston, K. (2018) How and why consciousness arises: Some considerations from physics and physiology. *Journal of Consciousness Studies*, *25*, 202–238.

# Decisions under uncertainty are more messy than they seem

Jarno Tuominen[a,b]

[a]Department of Psychology and Speech-Language Pathology, University of Turku, FI-20014 Turku, Finland and [b]Department of Sociology, University of Helsinki, FI-00100 Helsinki, Finland.
jarno.tuominen@utu.fi; jarno.tuominen@helsinki.fi
https://www.utu.fi/en/people/jarno-tuominen

doi:10.1017/S0140525X22002576, e109

**Abstract**

Conviction Narrative Theory (CNT) is conceptually so multifaceted as to make critical evaluation difficult. It also omits one course of action: Active engagement with the world. Parsing the developmental and mechanistic processes within CNT would allow for a rigorous research programme to put the account under test. I propose a unifying account based on active inference.

Behaviour under uncertainty is messier than proposed, and while laudable in its breadth, Conviction Narrative Theory (CNT) risks specificity for pluralism. Indeed, narratives themselves are presented in the target article as multifaceted, enveloping the functions of structure, simulation, affect-evaluation and communication. This fuzziness allows it to include a broad variety of cognitive functions and mechanisms, which while useful for more general descriptions, also makes its strengths and weaknesses as an explanation difficult to evaluate in detail. I propose that in addition to relying on narratives in situations of radical uncertainty we also have another option: Random exploration by acting on the world.

Consider, for example, an infant who in a situation of truly radical uncertainty (i.e., a situation with very shallow schemas, limited priors or disjointed narrative structure) acquaints with the world by putting things in their mouth, by wiggling and whacking the world. This is an example of model-free learning, where in novel situations we prod and poke to reveal the structure of the situation and possibly to alter it into a more familiar choice set. Only after we have inferred some structure to guide our predictions do we begin to form more complex model-based predictions to guide our actions (Castro-Rodrigues et al., 2022). Now it is not necessary that the inferred structure corresponds to any so-called ground truth, merely that it provides enough structure for action, that is, the reduction of uncertainty may be illusory. Alternatively, we may uncover this structure by asking others, of using others' representations as scaffolding for one's own.

Behind both attempts is the aim to reduce the uncertainty by acting on the world or by updating our internal model. Behind a variety of such uncertainty (or free energy) minimization accounts lies the free-energy principle (FEP, Friston, 2010), a broad formalization which arguably also engenders decision accounts from expected utility to softmax accounts (Friston et al., 2014). It proposes predictions to be hierarchically structured (Friston, 2010) with the task of the organism to reduce uncertainty thus maintaining the environmental fit of their "generative model." At the higher hierarchical end, predictions take the form of beliefs, with narratives as socially reinforced and integrated belief landscapes. These belief landscapes compete for valuation in a multidimensional matrix (for a main valuation account, see Sharot, Rollwage, Sunstein, & Fleming, 2022). A mismatch in the belief and the actual state of the world calls for one of two lines of action: Revising and updating the internal generative model; or active inference. Active inference refers to action selection as a process of imposing structure on the generative model organized from sensory data (Friston, Da Costa, Hafner, Hesp, & Parr, 2021). From this point of view, CNT would fit the FEP fold. The utility of such an account is singularly to reduce uncertainty in a fashion that can be approximated by Bayesian methods yet account for bounded rationality. FEP accounts sidestep the critique provided by the authors against Bayesian accounts that in situations of radical uncertainty priors cannot be set. Considering CNT a higher order method for uncertainty reduction, that is, as a way to increase the model fit of the environment and the generative model, the problem is removed: From the first-person view of the generative model, we can impose a model on the situation, even if its initial first-iteration fit were poor; to launch the competition for the best fit. After all, at some point, we stop dealing with uncertainty by putting things in our mouth for inquiry and utilize other methods.

CNT is praiseworthy in its scope and explanatory breadth. However, there are several avenues to begin clarifying these functions and the mechanisms. It is noteworthy that if were to replace the term "narrative" with that of "reflective consciousness" in the definition and describe it as:

> …structured, higher-order mental representations incorporating causal, temporal, analogical, and valence information about agents and events, which serve to explain data, imagine and evaluate possible futures, and motivate and support action over time (target article, Table 1 and sect. 5, para. 3)

> …we would not be that far removed from various accounts of consciousness overall. For example, biological realism considers our conscious experience as a spatiotemporal virtual reality, a simulation of the world beyond our senses (Revonsuo, 2006).

With such limitations in mind, I propose CNT would gain in explanatory depth by integrating a developmentally informed account of how we move from model-free learning to narratives under uncertainty. One could make the claim that only once the social complexity increases and our mental representations acquire complexity probabilistic reasoning becomes both limited and a resource strain – and when active engagement fails – do we switch to narratives, again unified under the function of uncertainty reduction and thus formalizable (for children faring well in probabilistic reasoning see Girotto & Gonzalez, 2008; Riggs, 2019).

A more thorough analysis would illuminate how the proposed mechanism functions: What is necessary and what is sufficient for CNT and its subcomponent dynamics. For example, it is suggested we simulate the structure of the mental representation of reality already before birth (Hobson, 2009), whereas narrative communication develops later as it requires theory-of-mind, and causal and temporal reasoning abilities (Stadler & Ward, 2005). Before such parsing, CNT is at risk of falling into the same mechanistic trap it sees as the demise of competing theories with the broader terms doing a lot of work.

**Financial support.** This research was written in the context of a research project that received funding from the Strategic Research Council at the Academy of Finland, grant number 335186.

**Competing interest.** None.

## References

Castro-Rodrigues, P., Akam, T., Snorasson, I., Camachi, M., Paixão, V., Maia, A., … & Oliveira-Maia, A. J. (2022). Explicit knowledge of task structure is a primary determinant of human model-based action. *Nature Human Behavior*, *6*, 1126–1141. https://doi.org/10.1038/s41562-022-01346-2

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, *11*(2), 127–138.

Friston, K., Da Costa, L., Hafner, D., Hesp, C., & Parr, T. (2021). Sophisticated inference. *Neural Computation*, *33*(3), 713–763.

Friston, K., Schwartenbeck, P., FitzGerald, T., Moutoussis, M., Behrens, T., & Dolan, R. J. (2014). The anatomy of choice: Dopamine and decision-making. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*(1655), 20130481.

Girotto, V., & Gonzalez, M. (2008). Children's understanding of posterior probability. *Cognition*, *106*(1), 325–344. https://doi.org/10.1016/j.cognition.2007.02.005
Hobson, J. A. (2009). REM sleep and dreaming: Towards a theory of protoconsciousness. *Nature Reviews Neuroscience*, *10*(11), 803–813.
Revonsuo, A. (2006). *Inner presence: Consciousness as a biological phenomenon*. MIT Press.
Riggs, A. E. (2019). Social statistics: Children use statistical reasoning to guide their inferences about the scope of social behavior. *Developmental Psychology*, *55*(1), 66–79. https://doi.org/10.1037/dev0000618
Sharot, T., Rollwage, M., Sunstein, C. R., & Fleming, S. M. (2022). Why and when beliefs change. *Perspectives on Psychological Science*, *18*(1), 142–151. https://doi.org/10.1177/17456916221082967
Stadler, M. A., & Ward, G. C. (2005). Supporting the narrative development of young children. *Early Childhood Education Journal*, *33*(2), 73–80.

# Authors' Response

## Narratives, probabilities, and the currency of thought

Samuel G. B. Johnson[a,b,c,d], Avri Bilovich[b] and David Tuckett[b,e]

[a]Department of Psychology, University of Warwick, Coventry CV4 7AL, UK; [b]Centre for the Study of Decision-Making Uncertainty, University College London, London W1CE 6BT, UK; [c]University of Bath School of Management, Bath BA2 7AY, UK; [d]Department of Psychology, University of Waterloo, Waterloo, ON N2L 3G1, Canada and [e]Blavatnik School of Government, University of Oxford, Oxford OX2 6GG, UK.
sgbjohnson@gmail.com
a.bilovich@ucl.ac.uk
d.tuckett@ucl.ac.uk

doi:10.1017/S0140525X22003223, e110

**Abstract**

Whereas most commentators agree about the centrality of narratives in decision-making, the commentaries revealed little consensus about the nature of radical uncertainty. Here we consider thirteen objections to our views, including our characterization of the uncertain decision environment and associated cognitive, affective, and social processes. We conclude that under radical uncertainty, narratives rather than probabilities are the currency of thought.

### R1. Three theorists in search of a brawl

> *Midnight. A full moon illuminates the town square, as a desperate trio vaults onto the scene. They pause, huffing and puffing as their academic lungs reach beyond their intended capacity, their sweat glistening. Torches. Pitchforks. Crazed villagers scurry from all directions: Your heroes are surrounded. Fifty-four manic eyes stare them down…*

A scene much like this was on our minds when we published this paper. Many of the proposed ideas confront not only neoclassical economics, but also most behavioral approaches, including much of judgment and decision-making research. Formal decision theory, artificial intelligence research, and standard economics are all deeply uncomfortable with the idea that there *could* often be situations requiring action in which there is no right answer, where probabilities are a poor guide. While behavioral economics rightly rejects the idea that we are optimal probabilistic decision-makers, it retains optimal probabilistic reasoning as a normative standard. We challenge this perspective too. Under radical uncertainty, we are not "good" or "bad" probabilistic actors. Probability simply has nothing to do with it.

We were looking for trouble.

Yet, in the town square that is this journal, we were confronted not by thugs, but by twenty-seven sets of experts in a dazzling array of fields, including cognitive psychology, neuroscience, philosophy, psychiatry, management studies, kinesiology, social psychology, affective science, Earth science, experimental humanities, semiotics, literature, organizational behavior, religious studies, and – the dismal science itself – economics. We were delighted to see commentaries from practitioners and applications of Conviction Narrative Theory (CNT) to issues in clinical practice (**Linkovski & Eitan**; **Siegel**), education (**Sheskin, Bogucki, Perry, & McAllister** [**Sheskin et al.**]), and society (**Campbell & Fonagy**). Pleasingly, not one of the commentators proposed immolation, defenestration, drawing-and-quartering, or even a good whipping. (Well, maybe a lash or two.)

As far as we can tell, *every* commentator agrees with our basic premise that narrative thinking is crucial to human decision-making. But the commentaries reveal a deep divide in the field over the relevance and even existence of radical uncertainty. Some agree that we need credible theories of how to take action under radical uncertainty. Others are more inclined to reduce radical uncertainty to risk – a position which we believe has limited applicability to the kind of real-life decisions that really matter as opposed to "small worlds" contexts. We doubt we can do much to shake these theoretical commitments – as **Grossmann, Meyers, and Eibach (Grossmann et al.)** suggest, people may have strong priors on the existence of radical uncertainty – but will do what we can to taxonomize the disagreements.

### R2. The problem

Real-world decisions differ in many ways from those studied in laboratories and textbooks. We must often make massively consequential choices when data are incomplete, the options are ambiguous, the future may not resemble the past, and the axioms of standard decision theory (see **Ford & Kay**) are not satisfied. Such decisions often require both commitment and flexibility over time. Above all, such decision-making is about taking action in the face of potentially paralyzing uncertainty. Choices like monetary gambles and perhaps even that facing our widow – who has been put through such an ordeal between our target article and interlocutors – may be amenable to standard analyses. But it is far less clear how career choices, climate change mitigation, pandemic preparedness, and economic development policy can be effectively understood this way.

While many of the problems associated with such everyday decisions differ from those of textbook "small-world" choices, two central issues must, as a matter of logic, be solved by any theory of decision-making. A central idea in cognitive science is that we do not act on the world directly, but through our mental *representations* of that external world (Fodor & Pylyshyn, 1981). These representations or beliefs need to be stored in a format that can act as a "currency of thought" that mediates between the information we pick up from the external world and our decisions to act on that world (the "mediation problem"; target article, Figure 1). In principle, many types of representations could fit this bill – anything sufficiently isomorphic to the world. But our decisions are determined not only by our beliefs, but by our values. We need some "driver of action" that combines beliefs and values into choices (the "combination problem"). In our view, the core logical challenge faced by behavioral decision theory is identifying cognitive architecture that can solve both problems simultaneously.

One important approach is to assign probabilities to various possibilities (our beliefs), utilities to each of those possibilities (our values), and choose the option that yields the greatest expected utility (our choice). In our terminology, the currency of thought is probability (our beliefs are represented in a way that both summarizes the world and feeds into our choices) and the driver of action is utility maximization (our choices reflect the weighted combination of beliefs and values). Determining those probabilities and utilities may be complicated, but the central idea is remarkably simple.

This style of theorizing has often been attributed to economics. But **Ford and Kay** note that formalizations of decision theory in terms of expected utility maximization (Savage, 1954) do not actually make *any* assumptions about psychology, so (i) such formal theories do not strictly treat probabilities as a "currency of thought" or utility maximization as a "driver of action," and (ii) CNT is not inherently in conflict with these theories. Yet this has not stopped economists, psychologists, and many policy-makers from treating these constructs in exactly the way we describe – tacitly taking expected utility maximization as the psychological basis of "rational" economic behavior. As one example among many, Becker (1968) writes of his economic theory of crime that "The approach taken here follows the economists' usual analysis of choice and assumes that a person commits an offense if the expected utility to him exceeds the utility he could get by using his time and other resources at other activities" (p. 176).

Behavioral economists have criticized this approach at length because people often assess probabilities wrongly. Fair enough. But we are pointing to a different and deeper problem: Often, decisions have to be taken – for instance whether to invest large sums in a new vaccine or to take costly steps to mitigate climate change – where *there are no probabilities to be wrong about*. In a world of risk, where outcomes are enumerable and probabilities calculable, expected utility maximization may make sense as a normative or even descriptive theory. In a world of uncertainty, where there is no way to assign probabilities even in principle, expected utility maximization not only is the wrong descriptive account, but is unintelligible as a normative theory. For these situations, behavioral economics inherits this unintelligibility by accepting the normative premise that uncertainty can be reduced to risk even while denying the descriptive claim that people do so optimally. The problem with using expected utility theory to explain decision under radical uncertainty is not that people are *bad* at probabilistic reasoning. The problem is that, under radical uncertainty, people *couldn't* be good at it.

We need a theory that accounts for action *with no probabilities at all*. This is where Conviction Narrative Theory comes in.

### *R2.1. Objection 1: Such a theory may not be possible*

Some commentators took issue, less with CNT's specific claims, but with its scope. Not to put too fine a point on it, **Newell and Szollosi** muse whether CNT is "a theory of everything or nothing." They see a tension between our aspiration to explicate the "currency of thought" and our claim to be developing a theory not *of* causation, analogy, emotion, etc., but of how these component processes *fit together* to account for decision-making under uncertainty. We do not see the contradiction. CNT says that, for decisions under radical uncertainty (though not necessarily in other contexts), narratives are the currency of thought. We go into considerable detail about what a narrative is and what processes it participates in. But our account sits at a higher level of analysis than theories of causation, analogy, and so on. There are multiple competing perspectives on these lower-level representations and processes, and we mostly avoid these debates because they are not part of CNT's theoretical commitments.

As inventions go, CNT is more like a car than a carburetor. The automobile's innovation was the particular way that the components are connected to one another, not the components themselves, such as wheels, brakes, or carburetors. A car, like a theory, is a tool for solving a particular problem. It is not a strike against its inventor that it includes parts invented by others, nor does the inventor deserve credit for their inclusion. But the fact remains: They take us further than a horse.

**King** asks whether a general theory of decision-making under radical uncertainty is even *possible*. The answer is probably "no" if King seeks a model, like expected utility theory, that can generate choices from some tractable set of inputs (e.g., subjective probabilities and utility functions). We doubt that even a descriptive, much less a normative, theory of this sort is possible. Descriptively, narratives are likely to depend on unique features of the situation and the decision-maker's idiosyncratic beliefs and values, which may be forbiddingly difficult to systematize. Normative theories of decision-making under radical uncertainty in general are likely to fall short for similar reasons to probabilistic approaches – who is to say what the right action is in a non-stationary world where the underlying model is not only unknown but changing?

But less ambitiously, we do believe an *explanatory* rather than predictive or normative theory is possible. We believe we can characterize the *kind* of mental representations and processes that are typically used to make such decisions, even if the *content* of those decisions is going to depend on too many unspecifiable, context-sensitive factors for a complete predictive theory. The more precisely the field can characterize those lower-level components (e.g., causation, analogy, affect), the more explanatory the resulting product (CNT + the best theories of causation, analogy, etc.) will be. We also believe that by focusing decision research on the conditions in which conviction (rather than a "correct" choice) is achieved, an agenda is created that is of much greater practical use.

## R3. Context

To restate CNT briefly: For many everyday decisions, people adopt *narratives* – mental representations coordinating causal, analogical, temporal, and valence structure – to explain decision-relevant information and link proposed actions to intended outcomes. Narratives are the joint product of individual cognition and the social environment: People pick up on "narrative fragments" from others which feel right in context and are incorporated into individual narratives to support action. Once a narrative is adopted that makes sense of the evidence, people use the causal structure in that narrative to simulate or imagine the future that would unfold if they chose a particular course of action. They then deploy the same emotional processes they use to appraise actually present situations to appraise the imagined future. If the appraisal yields approach emotions, then they choose that course of action; if it yields avoidance emotions, they choose a different course. Since decisions must manage doubts and eventualities over time, people can do so either adaptively in an "integrated state" that maintains commitment while being open to evidence and potentially course corrections, or maladaptively in a "divided state" that repels doubts by blindly clinging to a narrative until reality intervenes. People may chronically differ in their tendency toward either stance, but situational forces such as organizational structure and normative practices can shift people toward one or the other (Fenton O'Creevy & Tuckett, 2022).

We view this process not as infallible, but as by-and-large adaptive given the constraints posed by the environment under which many everyday decisions are taken. We describe four such constraints at the beginning of the target article: Such decisions are often taken under radical uncertainty (meaningful probabilities do not exist) and fuzzy evaluation (a precise and commensurable utility function cannot be characterized), the decision often must be sustained over time (they require a long-term commitment), and they are often socially embedded (information is processed not only through individual cognition, but through our interactions with others). Thus, we agree with **Marewski** that both blades of Simon's (1990) scissors are needed to understand decision-making – the external environment and internal processes.

Our point is that narratives can help to cope with all of these features of the decision environment while implementing the internal psychological processes (sense-making and imagination) that so often seem to characterize the phenomenology of decision-making. Narrative evaluation relies on heuristics rather than probabilities. Narrative simulation relies on affect rather than utilities. Narratives can resist change, permitting commitment (though also risking ossification, as **Breithaupt, Hicks, Hiskes, & Lagrange** [**Breithaupt et al.**] and **Campbell & Fonagy** rightly note). Narratives can incorporate information from the social environment, permitting the forces of cultural evolution to apply selection pressures to our beliefs alongside individual cognition. This socially distributed aspect of narrative decision-making is an important part of how people often make reasonably good decisions despite environmental and cognitive constraints.

Although we received minimal pushback on our points about fuzzy evaluation, commitment, or social embeddedness, we encountered several objections to our characterization of radical uncertainty.

### *R3.1. Objection 2: Radical uncertainty does not exist*

We sense a deep divide among our commentators about the very existence of radical uncertainty, which we believe to represent a disagreement across many fields more broadly. For example, **Jinich-Diamont and Christov-Moore** suggest that uncertainty falls on a continuum, where our probability distributions in more uncertain situations depend more on prior probabilities (generated from a higher-order probability distribution) than on new evidence. In fact, this is an eliminative account of radical uncertainty – it's probabilities all the way down. **Friston**'s commentary can also be read this way. Conversely, many others seem to view the issue as decided in the opposite direction, including **Ford & Kay**, **Gigerenzer**, and **King**. **Marewski** also notes Simon and Newell's distinction between well-structured and ill-structured problems, which we see as closely related, since one way a situation can be uncertain is when the outcomes cannot be enumerated. Although most commentators left their view on this topic implicit – and we are reluctant to take words written between the lines and place them firmly in people's mouths – we suspect many commentators and readers find themselves attracted more to one or the other of these positions.

What we are *not* saying is that, under radical uncertainty, every possibility is equally plausible and anything goes, but rather that probabilities are incalculable. For example, here is Keynes' (1937) famous passage:

> By "uncertain" knowledge, let me explain, I do not mean merely to distinguish what is known for certain from what is only probable. The game of roulette is not subject, in this sense, to uncertainty…The sense in which I am using the term is that in which the prospect of a European war is uncertain, or the price of copper and the rate of interest twenty years hence, or the obsolescence of a new invention, or the position of private wealth-owners in the social system in 1970. About these matters there is no scientific basis on which to form any calculable probability whatever. We simply do not know.

We can read **Friston** and **Jinich-Diamont and Christov-Moore** in two ways. On the one hand, coming from the relatively stable world of biological functioning (e.g., vision or motor control), we can suppose they are mainly thinking about contexts which change very slowly and are strongly path-dependent. It means they can usually discount the significance of the kind of rapid and unexpected innovation that characterizes the social and economic phenomena with which Keynes was preoccupied. In this case, our differences with them are ones of focus and framing. On the other, if not and they really have in mind decisions about pandemics, climate change, economic policy, etc., then their approach, in which decision-makers rely on prior probabilities even under radical uncertainty, is hard to support. We doubt that people have prior probability distributions over such events as those Keynes is describing. Perhaps, when pressed, people might provide such numbers. But such estimates could only be improvised in the moment and have little relationship to behavior (here, we suspect we are aligned with **Chater**).

We accept we will not convince everyone with our negative case against probabilistic approaches. Preexisting philosophical and disciplinary commitments preclude this. Yet, in the target article, we also marshal many sources of positive evidence for narrative influences in decision-making. Some of this evidence is potentially reconcilable with probabilistic approaches to narrative, while some would be very difficult to reconcile (e.g., the "digitization" results in Johnson, Merchant, & Keil, 2020, as **Greco** notes). For **Friston**, we suspect his proposed probabilization of narrative decision-making would be a satisfying theory that could formalize narratives as a tool for decision-making under risk (we take it that for Friston, all decisions are risky in the sense that they can be formalized using probability theory). We are much less bullish on this approach because we reject the rejection of radical uncertainty. This leads naturally to the next objection.

### *R3.2. Objection 3: CNT [does] [does not] explain decisions under risk*

CNT has little to say about decision-making under risk. Probabilistic theories, such as classical decision theory or Bayesian models or heuristics-and-biases models that take probabilities as inputs, are sensible contenders and we do not attempt to adjudicate this debate. **Beach and Wise** believe this is an

https://doi.org/10.1017/S0140525X22001157 Published online by Cambridge University Press

unnecessary limitation. We see it as an empirical question, as either answer could be justified theoretically.

If so many decisions are taken under radical uncertainty that our cognitive architecture is primarily attuned to such contexts, then plausibly narrative cognition applies to risky decisions as well. One way that this could be true is if we think of some narratives as *licensing* probabilistic reasoning – a point closely related to **Grossmann et al.**'s suggestion that radical uncertainty is itself a narrative and to **Gigerenzer**'s observation that the idea of a "small world" is a compelling narrative. To use Beckert and Bronk's (2018) term, "calculative devices" can be embedded in a narrative, as when a gambler decides to adopt a particular mathematical procedure as an input to their decision-making strategy. Less sanguinely, policymakers may reify a particular economic model as literally true, applying its mathematical formalism to a situation in which it is inappropriate. Orthogonal to whether the calculative device is *appropriate* to a situation, its mathematical application may itself be *deployed* correctly (as when seasoned weather forecasters apply their models) or incorrectly (as in the many mistakes identified in judgment and decision-making research).

Alternatively, since probabilistic reasoning is indeed a valuable way to decide under risk, we do not rule out the possibility that we have separate mechanisms for doing so. Indeed, CNT is not trying to explain *all* decisions. Some commentaries point to very low-level decisions that we agree may not be mediated by narratives. **Cañal-Bruland and Raab** note that many motor decisions, such as where and when to hit a baseball, do not seem to be especially cognitively mediated. **Andreetta, Spalla, and Treves (Andreetta et al.)** look to decisions such as selecting the correct answer in a recognition memory task, questioning whether such decisions are affectively mediated. We don't know; we simply see this as a non-overlapping magisterium. Problems such as when and where to hit a baseball are far more constrained and structured than the sorts of problems CNT is intended to address.

If we do have separate mechanisms for dealing with risk versus uncertainty, then as **Newell and Szollosi** note, we would need to have some way to *identify* such situations. Alternatively, if even risky decision-making is mediated through narratives, we still need principles for identifying when to apply narratives that include calculative devices. **Grossmann et al.** make an impressive array of intriguing suggestions in this connection, including the role of folk theories, individual differences, moral norms, motivation, and culture.

## R4. Representations

We characterize narratives as "higher-order" representations; they do not necessarily confine themselves to causal structure, but can coordinate causal structure with analogical, temporal, and valence information. We believe that this kind of representational flexibility is integral to the sophistication of human thought. Despite its importance, there is little theoretical or empirical work examining how these different kinds of lower-level representations are coordinated. We sketched how these representations might look, but we readily acknowledge that much more needs to be done. Here, we consider three challenges to our sketch.

### *R4.1. Objection 4: CNT's representational framework is too complicated*

**Newell and Szollosi** recommend that, in lieu of narratives, we focus on simpler representations. They give as examples the study of how people represent integers and event frequency.

We certainly have no objection to research that takes a more minimalist or even reductionist approach to understanding mental representations or processes. Such strategies have been massively successful in cognitive science for understanding countless issues. But this does not entail that they are the *only* legitimate scientific approach or that they are well-suited to solving *every problem*. Indeed, Newell and Szollosi unfavorably juxtapose our view that qualitative methods are superior to experiments for understanding the *prevalence* of narrative thinking with our view that experiments are superior to qualitative methods for understanding the *processes* of narrative thinking. But this is not a contradiction! Different tools are suitable for different purposes. Lab experiments risk creating thoughts or behaviors in artificial settings, where they would not exist in a naturalistic setting – so they are not ideal for understanding how frequent a behavior is in the real world. Interviews can suffer from post-hoc rationalization and the opacity of introspection – so they may not be ideal for understanding some psychological processes. Methodological imperialism is equally myopic whether it is staked by psychologists on behalf of lab experiments, behavioral economists on behalf of field experiments, economic theorists on behalf of formal models, or social scientists on behalf of interviews and ethnographies. Complex problems require all the help we can muster if we are ever to triangulate the truth.

As for the specific suggestion to build up from simple to complex representations, we believe success is far more likely if we approach the problem from both angles. After all, there are *many* investigations of simpler representations and processes, including much of our own work and arguably *most* of cognitive psychology. We would channel here (a different) Newell (1973), who argues that "you can't play 20 questions with nature and win." The elder Newell means that we need to conjoin experimental evidence that dissects component processes ("playing 20 questions") with integrative models that explain how these processes work in concert to yield thought and behavior. In the case of CNT, this is not (yet) developed to the level of a mathematical model – and, indeed, part of the reason is that cognitive science does not yet have complete models of the component processes. Yet we believe that theorizing the *relationships* between cognitive processes must take an equal role to examining how those component processes work. We need both cars and carburetors.

### *R4.2. Objection 5: CNT's representational framework is too simplistic*

Coming from the totally opposite end, **Caldwell** suggests that our representational model of narratives is too simple. We do not intend for our sketch of a representational theory of narratives to be the final word on the subject, and we agree with Caldwell that it can be developed further. Some of Caldwell's extensions are plausible and should be considered in future work. Yet extensions of this framework should be (i) mindful of the empirical evidence, and (ii) careful to preserve the features that make it suitable for dealing with action taken under radical uncertainty.

As an example of the first issue, Caldwell suggests that people represent some causal relationships as stronger than others. This may be true, but the point is controversial in the causal cognition literature (e.g., Yin & Sun, 2021) and we are ourselves unsure.

As an example of the second issue, we suggest caution in installing "a scalar affective value" into each node. We are unsure that this is an advance over representing utilities. As Walasek and Brown (in press) forcefully demonstrate, the problem with utilities is shoving preferences into a one-dimensional common currency, as a scalar affective value would. CNT abandons this approach in favor of affect because affect is multidimensional, motivates action, and is an evolved system for adaptively dealing with real situations. As no small bonus, emotions are known to exist, whereas no one has ever actually seen a utility function in the wild (though we understand that there is a large bounty for one caught alive).

### *R4.3. Objection 6: Narratives are too incoherent to be the building blocks of knowledge*

**Chater** asks whether narratives are stored in long-term memory and applied off-the-shelf to situations as they arise, or instead improvised in the moment and constructed on the fly. On one extreme, Chater suspects that narratives are constructed as post-hoc justifications. On the other extreme, **Caldwell** suggests that all of our knowledge is constructed in one grand narrative from which we extract a subset at any given time – a point **Beach and Wise** echo in their invocation of a "prime narrative" in the Theory of Narrative Thought.

While we agree with **Chater** that narratives can be inconsistent, this is not a sufficient basis to reject the idea that we store narratives that support action. Narratives are beliefs, and although Chater (2018) believes that "no one, at any point in human history, has ever been guided by inner beliefs or desires," we do not share this belief. (Unless Chater is right, in which case neither of us have beliefs to disagree about!) Most beliefs are much more pedestrian than unconscious desires or repressed memories: Bees can sting, fertilizer makes crops grow, my boss is grumpy on Mondays, J-walking is taboo. Such beliefs are not consciously accessible most of the time, but plainly guide our decisions; they are not post-hoc confabulations for avoiding the beehive, purchasing fertilizer, complimenting the boss, or waiting for the traffic signal. The most extreme version of Chater's view appears untenable: The building blocks of narratives – including causal and analogical knowledge, social norms, and probably simple intuitive theories – seem clearly to be stored as beliefs in long-term memory.

Yet we would not go as far as **Caldwell** or **Beach and Wise** in the opposite direction. Caldwell suggests that narratives are excerpted as subgraphs from a "causal graph…representing the agent's whole mental model of the world." One needn't embrace a Chaterian view of mental flatness to find this implausible. It is unlikely that people could maintain a consistent model of this size in their minds, both for theoretical reasons (e.g., the "frame problem" mentioned by commentators) and empirical reasons mentioned by Chater. Indeed, our own prior work suggests that causal representations are remarkably discrete (Johnson & Ahn, 2015, 2017). People believe that there is a causal relationship, for instance, between wine (A) and sleep (B) and between sleep (B) and dreaming (C), but not between wine (A) and dreaming (C). People have separate, schematized causal mechanisms for the soporifics of wine and for the phenomenology of dreams. Causation is mentally represented in "islands" rather than "networks."

Rather than impaling ourselves on the horns of this dilemma, we tentatively offer the following resolution. We agree with **Chater** that narratives are not the "building blocks" of knowledge. That distinction goes to the lower-level representations that are *coordinated* by narratives (e.g., causal and analogical schemata). But those lower-level schemata must often be fleshed out in more complex narratives to support sense-making, prediction, and action. Sometimes this fleshing out happens slowly, through cultural evolution and social development, whereas other times it may indeed occur on-the-fly as Chater suggests, to fit a particular situation.

Unlike **Chater**, however, we believe that narratives can also be stored, retrieved, and elaborated over time. This is central to CNT's account of why decision-making can be reasonably adaptive, even under radical uncertainty. Sustained representations of narratives enable two feedback loops. First, they allow feedback from our actions to impact our narratives on an individual level. Decision-makers potentially revise narratives in light of whether actions taken on their behalf work out well or not. Second, sustained individual representations of narratives underlie the shared narratives that are shaped by cultural evolution. Entire societies and subcultures can fall under the spell of a shared narrative, for good or for ill, as illustrated by examples from **Campbell and Fonagy**, **Gigerenzer**, and **Lightner**. This depends on shared representations, such as transactive or collective memory (Hirst, Yamashiro, & Coman, 2018), which conjoins memories stored in individual minds with social processes.

Thus, we reject the claim that narratives are merely epiphenomenal constructions after the fact, made to justify our decisions. It is certainly true that narratives do often play a justificatory role (Cushman, 2020; Mills, 1940), as we note in the target article. But we think it clear that they can also cause decisions. For one, the anticipated need to justify a behavior can *itself* change our behavior; as Mills (1940) puts it, our justifications must make sense to collaborators, requiring that they fit a "vocabulary of motive." But even beyond their justificatory role, causal narratives seem to cause behavior. Surely causal cognition functions to allow successful interventions on the world (Woodward, 2003) and analogical reasoning functions to apply lessons from one situation to another, supporting action (Holyoak, 1985). Many experiments cited in the target article show how altering one aspect of a mental representation (causal, analogical, temporal, or valence information) impacts downstream decision-making. Narratives can serve a justificatory function, and these communicated narratives may even differ from the narratives that drove behavior. But recognizing that narratives can be constructed *post hoc* does not mean that a decision was not also driven by a narrative *ex ante*.

## R5. Explanation

Turning from representations to processes, we first highlight the role of narratives in making sense of decision-relevant information by imposing structure. The target article made two claims about this process. First, it relies on heuristics; second, explanatory fit is experienced affectively (some narratives "feel right"). Many of the more critical commentaries highlighted concerns about this process.

### *R5.1. Objection 7: CNT does not provide an account of how narratives are generated*

Several commentaries point out that we do not provide an account of how narratives are generated, only how they are evaluated once they are generated (**Caldwell**, **Greco**, **Newell & Szollosi**). This is not an issue that CNT considers in detail, although it is indeed

https://doi.org/10.1017/S0140525X22001157 Published online by Cambridge University Press

important. As Greco notes, solving this issue is an essential part of the transition from a "grand world" problem to a more tractable "small-world" problem. For instance, **Friston** seems to equate the problem of Bayesian model selection with the problem of hypothesis generation, but such a process requires candidate models to be considered. Some process that sits outside of probabilistic reasoning would be necessary to generate the structures to be selected, even if the selection process operates along Bayesian lines.

Much of the work of narrative generation is in fact outsourced to the social, historical, and family environment. Certainly, new "grand narratives" of the sort described by **Gigerenzer** and **Lightner**, or those discussed by Shiller (2019), arise only once in a great while, emerging slowly through cultural evolution. (In this sense, the question "how are narratives generated?" is rather like "how are species generated?") On a much smaller scale, though, people likely combine causal and analogical knowledge in new ways in everyday life, such as when we consider a new career or relationship. CNT does not provide a detailed account of this process, but we do believe that a narrative approach helps to focus attention in the right direction.

First, narrative accounts can draw on the broader literature on how people generate causal hypotheses, including through heuristics. The central idea is that effective generation strategies (i) rely on easily accessible information and (ii) cue hypotheses that have a reasonable chance of being relevant and correct. Many of these strategies are related to time (e.g., Lagnado, Waldmann, Hagmayer, & Sloman, 2017; Rottman & Keil, 2012), which is one reason to think narratives coordinate temporal information. One example, briefly mentioned in the target article, is our own prior work on how event structure is used to narrow the space of potential causes for a specific event (Johnson & Keil, 2014). Research on event perception has found that people segment experience into discrete events (Zacks & Tversky, 2001) at multiple levels of granularity (i.e., higher-order events that subsume lower-level events, such as a trip to the mall that includes episodes for visits to particular stores). Johnson and Keil (2014) found that people use two heuristics derived from event structure for narrowing candidate causes: (i) Events at one level of granularity are preferentially considered as candidate causes (i.e., matching higher-level effects with higher-level causes); and (ii) for events at a lower level of the hierarchy, other lower-level events that are part of the same higher-level event are preferentially considered as candidate causes.

Second, hypothesis generation must draw not only on information readily available from the immediate environment, but from long-term memory. Thus, a detailed understanding of how the lower-level building blocks of narratives are represented and indexed in memory will be crucial. When narratives must be generated bottom-up, the generation process presumably begins with the evidence to be explained, which acts as retrieval cues for general knowledge that may be incorporated into the narrative. If the evidence includes observations, A, B, and C, then the generation process will involve memory search for causes and analogies related to A, B, and C. Indeed, we know that causal relationships act as stronger retrieval cues than non-causal relationships with equally strong associative strength (Fenker, Waldmann, & Holyoak, 2005). Plausibly, the narrative would be built up until it is deemed a sufficiently complete explanation (Korman & Khemlani, 2020).

Third, notice that storing narratives in memory greatly simplifies this process, as does adopting shared narratives from the social environment. Rather than carrying out a potentially arduous process of narrative construction, one could simply pull a narrative off the shelf when triggered by a given situation, perhaps with some tailoring; or multiple narratives might be retrieved and then evaluated. This is one additional reason to be skeptical of the idea, à la Chater, that narratives are always or usually constructed on the fly.

### R5.2. Objection 8: CNT is not specific enough about what heuristics are used

**Greco** mentioned that we were not very specific in the target article about what heuristics are used. While CNT is itself agnostic on this question, there is a growing literature on exactly this issue that we reference in the target article (e.g., Horne, Muradoglu, & Cimpian, 2019; Johnson, Rajeev-Kumar, & Keil, 2016; Lombrozo, 2016). As one example among many, people use simplicity as a cue to an explanation's prior probability and complexity as a cue to its likelihood or fit to the data (Johnson, Jin, & Keil, 2014; Johnson, Valenti, & Keil, 2019b; Lombrozo, 2007), circumventing the need for exact probabilistic computations. This literature mostly looks at much simpler kinds of representations (of the sort that **Newell & Szollosi** favor studying), so future research should generalize these heuristics to more complex narrative structures, understand how they interact, and study how they are prioritized when in conflict.

Some of the commentators mentioned still other heuristics. **Gigerenzer** brings up the fascinating point that narratives can themselves select heuristics (his Protestant Work Ethic example). **Schwarz** highlights the importance of metacognitive cues, such as fluency, with a related point made by **Jinich-Diamant and Christov-Moore** about "interoception" (observation of internal states). We agree that such cues are important. As Schwarz notes, fluency is closely related to our point that we select narratives that "feel right," and thus to the construct of explanatory "satisfaction" studied in the explanation literature. The causes of explanatory satisfaction are often, but not always, structural features such as simplicity (Lombrozo, 2007) and explanatory scope (Khemlani et al., 2011). That is, feelings are often the proximate cause of behavior, even as they implement intelligent and context-sensitive heuristics or strategies for selecting actions. We suspect that Schwarz is sympathetic to this point, although he also discusses evidence that incidental emotions can influence fluency for reasons unrelated to structure.

### R5.3. Objection 9: Heuristics could be used to estimate probabilities rather than narratives

There are three versions of this objection.

First, **Greco** argues that probabilistic theories have an equal right to invoke heuristics as CNT does. Thus, evidence that people use heuristics to evaluate explanations should not count as evidence for CNT. We agree with this in principle, but in practice probabilistic theories of cognition tend to reduce heuristics to *approximation* techniques for performing probabilistic calculations. For example, Griffiths, Chater, Norris, and Pouget (2012) explicitly contrast their view of heuristics that we "can connect directly to optimal solutions" (p. 417) against "bag of tricks" views. This is a plausible move for probabilistic accounts, which link to more recent developments in "resource rationality" (Lieder & Griffiths, 2020), and which may be successful for modeling cognition under risk. We remain skeptical of "optimal" solutions, whether exact or approximate, under radical uncertainty.

Second, **Jinich-Diamant and Christov-Moore** argue that CNT actually is sneaking in probabilities under the cover of darkness when it invokes heuristics such as "simpler explanations are *likelier* than complex ones." This is a mistake. A heuristic that compares two possibilities and concludes that one is likelier does not logically necessitate probabilities, but only a comparative (rather than absolute) judgment of likelihood. Radical uncertainty is not the same as saying that anything goes. Some possibilities (there will be another pandemic or development of cheap nuclear fusion) are more plausible than others (there will be a civilization-ending nuclear war or development of faster-than-light travel) even if we cannot assign probabilities to them. This is true even without radical uncertainty. If you see one bag with many black marbles and another with only a few, your approximate number system (Feigenson, Dehaene, & Spelke, 2004) can estimate which bag has more marbles in it. You would conclude it is likelier that a marble drawn from one bag is likelier to be black than one drawn from the other, even if you have little idea how many marbles are in each bag. Crucially, while such behavior could be rationalized probabilistically from the outside, probabilities themselves need not be *mentally represented* in making this judgment. To view it as impossible to deem one event more likely than another without representing probabilities is to trivialize probabilistic approaches.

Third, a possibility not raised by commentators but which troubles us nonetheless, is that even if probabilistic theories could not be appropriate *normative* accounts under radical uncertainty, people might nonetheless *impose* probabilities where they ought not. On the one hand, much of the evidence reported in the target article contradicts this perspective for simple everyday decisions under radical uncertainty. But of course, inappropriate probability-based decisions can be observed, as when value at-risk models made disastrous probabilistic assumptions in the lead-up to the 2008 financial crisis, ignoring the possibility of things never previously encountered (Taleb, 2007). In fact, we suspect that the tendency to conflate risk with uncertainty is primarily a (recent) cultural phenomenon that has occurred as the proliferation of risk management tools and other quantitative models have appeared to offer greater levels of "authoritative" precision. This is one way that "the belief that we live in a small world is itself a powerful narrative," as **Gigerenzer** put it.

### *R5.4. Objection 10: CNT can be subsumed into probabilistic theories*

Several commentaries compare CNT to probabilistic approaches, such as Free Energy Theory, including **Friston**, **Jinich-Diamant and Christov-Moore**, **Solms**, and **Tuominen**. Insofar as these approaches are committed to probabilistic *representations*, we view this reduction as implausible under radical uncertainty (see Objection 1). For example, **Jinich-Diamant and Christov-Moore** suggest that we can "take narrative to be the structure of higher-level priors in the nested predictive hierarchy" and feelings as "low dimensional representations of interoceptive predictions about the expected consequences of visceromotor commands on the body's internal milieu." We are unsure how literally this is to be taken, but, needless to say, we are wary of such representational claims.

More broadly, we think it crucial to be clear about what level of analysis a theory is operating at. This can help to clear up confusion. For instance, sometimes the claim is made (e.g., by **Jinich-Diamant & Christov-Moore**) that probabilistic theories do not view our minds as *actually* computing probabilities, but that probability distributions can be constructed that *approximate* what the mind is doing. This is, roughly, the move from Marr's (1982) algorithmic to computational level. We have no problem with this as a research strategy (and said so in sect. 4.1 of the target article), with three caveats.

The first is that such theories should make falsifiable predictions which should be reasonably close matches to the empirical evidence. Some of the empirical results described in the target article are flatly inconsistent with probabilistic theories, with work on "digitization" (Johnson et al., 2020; Murphy & Ross, 1994) striking at fundamental claims about probabilistic thought. This is not to say that theoretical reconciliations could not be generated, but that we have not yet seen attempts to do so.

The second is that we must be clear on their theoretical status (i.e., what scientific problem they are solving). Computational-level theories are useful for characterizing why the outputs of a cognitive system are intelligent given the problem and constraints, but they explicitly are *not* mechanistic theories. The reason is obvious: If the theory is not saying that people represent probabilities in their minds but instead are doing something well-approximated by Bayesian inference, a mechanistically explicit account would need to say *what that thing is*. We are not claiming that CNT is a fully mechanistically explicit theory in the sense that in its current state it could be simulated on a computer – but it *is* mechanistically explicit in the sense that it claims that people *do* represent narratives in their minds, and *do* use them in processes of explanation, simulation, affective evaluation, and communication. There is much more to be said about all of these things, but CNT does specify what representations and processes underlie decisions under radical uncertainty.

The third is that computational-level accounts often depend on a "rational analysis" of a task (Anderson, 1990; Chater & Oaksford, 1999). We agree that this is valuable. However, we are skeptical that a probabilistic analysis of choices under radical uncertainty *is* a rational analysis. Approaches such as Free Energy Theory seem potentially well-suited for understanding evolved biological functions in stationary environments where probability distributions are learnable because they do not shift; under such conditions, the theory's commitment to optimization (Friston, 2010) can be appropriate. Yet, while **Friston** notes that we can be uncertain about hidden states, model parameters, or models themselves, he does not consider that the underlying model *itself* can often shift in complex decision environments such as the economy – they are non-stationary. The assumptions required to "rationally analyze" radical uncertainty using probabilities in fact *assume away radical uncertainty*.

Although the four commentaries written from an explicitly probabilistic perspective all are couched in terms of Free Energy Theory, other probabilistic perspectives are also relevant. One family recently gaining traction is resource rationality (Lieder & Griffiths, 2020), which considers bounded rationality from a Bayesian perspective. These accounts invoke tools such as approximation algorithms and attempt to give Bayesian rationales for specific heuristics given cognitive constraints. We are sympathetic to such approaches, particularly to the sampling approaches that have recently begun to accrue empirical support. Such approaches can yield "Bayesian brains without probabilities" (Sanborn & Chater, 2016), a view which seems far more plausible under radical uncertainty compared to approaches that invoke explicit probabilistic representations. Examining whether sampling and narrative approaches can be unified would be an exciting theoretical task.

## R6. Simulation

With a narrative in mind to explain the past, CNT holds that decision-makers use the causal and temporal structure embedded

https://doi.org/10.1017/S0140525X22001157 Published online by Cambridge University Press

in the narrative to simulate the consequences of potential actions. These simulations occur one a time, with people focusing on one particular imagined future that feels accurate rather than weighting combinations of different futures.

### R6.1. Objection 11: Simulation changes narrative plausibility

**Enz and Tamir** describe some fascinating experiments showing that merely simulating the future changes its perceived likelihood, in part by making alternatives harder to retrieve. This mechanism is compatible with, but distinct from, those invoked by CNT. Given that a narrative is often compatible with multiple possible futures, a further account is needed of how imagined futures are generated from narratives, unifying with greater detail the evidence discussed in the target article. The path-dependence proposed by Enz and Tamir will likely be an important part of such a theory, especially for the important subset of decisions that are made repeatedly or sustained over time.

### R6.2. Objection 12: People can and should simulate multiple possibilities

**Breithaupt et al.** balk at our suggestion that people simulate a single future at a time. To be clear, we are not suggesting that people only *consider* a single future. Rather, we are saying that people consider one at a time, cannot consider them all, and adopt one (provisionally, in the case of an integrated state) to guide actions sustained over time. Thus, multiple futures can be considered and compared, with only one (e.g., the most plausible one) adopted as true. What we *do not* believe people can typically do is to contemplate multiple futures *simultaneously*, weighing predictions by the relative probability that each of those futures might happen. This is the point made by the "digitization" experiments discussed in the target article (Johnson et al., 2020; Murphy & Ross, 1994).

Yet even provisional *adoption* of a single imagined future can get one quite far. First, we need not always adopt the most plausible future, but instead sometimes select futures that highlight threats (an important point made by **Beach & Wise**). Indeed, the boundary conditions on "digitization" tend to feature exactly such situations, such as dangerous categories (Zhu & Murphy, 2013) or immoral people (Johnson, Murphy, Rodrigues, & Keil, 2019a). Second, despite *cognitively* adopting a single future as true (in the sense that they do not average over alternatives), people can still experience *metacognitive* uncertainty about that belief if the adopted future is not seen as much more plausible than alternatives. In Johnson et al. (2020), participants ignored less-likely alternatives when making predictions, but when more-likely and less-likely alternatives implied different futures, participants were less confident in their predictions. The adoption of a single future, coupled by metacognitive uncertainty that fuels epistemic humility, can support a committed course of action while maintaining an open stance toward alternatives (an "integrated state"), both in the sense of willingness to revise one's narrative as new information accrues and of "hedging" behavior like that described by **Breithaupt et al**. This strikes us as an important direction for future research.

## R7. Affective evaluation

With an imagined future in mind, the decision-maker uses the same affective system to appraise that future as they would use for an actually present situation, which in turn feed into our motivational system to approach or avoid choices that lead to those futures. While emotions start out as adaptive mechanisms to prioritize needs, through development they come to internalize social input and reflect cultural wisdom (Scherer, 2005).

We received little push-back on the specific role of emotions invoked by our theory, including our idea that appraisals could be carried out over "default" dimensions (those invoked in standard appraisal theories of emotion) as well as "ad hoc" dimensions (cued by specific goals) and our ideas about how emotions help to manage decisions that must be maintained over time.

### R7.1. Objection 13: CNT unduly prioritizes affect at the expense of cognition

Despite relatively little pushback on specifics pertaining to emotion, beyond what we mentioned in earlier responses, we did get the sense that some commentators were uneasy with the extent to which CNT prioritizes affect over cognition. We simply think affect versus cognition is a false dichotomy. Rather, CNT is a theory of how affect and cognition relate, support one another, and jointly support action. Affect guides which narratives we adopt (though not *only* affect, we hasten to add in response to **Schwarz**) and affect appraises potential choices via the imagined futures they invoke (though this need not be the only effect of imagination, as **Enz & Tamir** note). Moreover, in adaptive decision-making characterized by "integrated states," cognition and affect work together to maintain conviction and potentially revise narratives in light of new information. Affect and cognition are essential and inseparable in adaptive decision-making, with neither more important than the other.

## R8. Communication

CNT encompasses both cognitive and social processes. Fragments of (individual) narratives are communicated within social groups, giving rise to shared narratives that in turn feed back into individual narratives. This is important both in solving the challenging problem of narrative generation (as **Greco** and several other commentators note) and in harnessing the adaptive power of cultural evolution to explain how people can make choices that mostly work, in a world of uncertainty. Narratives facilitate social coordination (via reputation-tracking and persuasion), drive social learning, spread through social networks, and evolve. **Pelletier, McLaughlin, and Boespflug (Pelletier et al.)** highlight the usefulness of a semiotic approach for understanding these functions, while **Jones and Hilde-Jones** extend some of these ideas more formally, identifying reasons why narratives act as "cultural attractors" that might be resistant to change (for better and for worse).

Other commentators highlight both the positive (**Lightner**) and negative (**Campbell & Fonagy**) implications of narratives' privileged status in communication. Together with the commentaries by **Linkovski and Eitan** and **Siegel** applying CNT to clinical practice, **Campbell and Fonagy**'s commentary suggests the intriguing possibility that individual psychopathy is often linked to maladaptive (individual) narratives, while "collective psychopathy" is often linked to maladaptive shared narratives. Given the pressing importance of issues such as polarization and misinformation, studying such issues through a CNT lens may be valuable.

### R8.1. Objection 14: There is no guarantee that cultural evolution will select adaptive narratives

This objection was actually not raised by any of the commentators, but we think we need to make clear it is an important issue for future research.

Consider **Lightner**'s interesting commentary on religious narratives. Lightner asks why narratives are so often false, concluding that even false narratives can enhance fitness by simplifying problems, solving long-term collective action dilemmas, and harnessing powerful emotions to generate conviction to act and coordinate behavior. Whether such narratives are actually true has little to do with whether they are adaptive. This account is quite consonant with CNT.

Presumably, the selection story here is a group selection one – groups that evolve such narratives are likelier to survive and expand, and thus the narrative spreads with the group. But an important point raised by Dawkins (1976) is that "memes" such as narratives, like genes, are selfish. That is, selection favors memes that themselves survive and propagate, not necessarily memes that help their hosts to survive. These *can* overlap, as Lightner argues for religious narratives, but they need not. The examples of misinformation raised by **Campbell and Fonagy** may fall into this category.

Thus, what guarantee (if any) is there that cultural evolution of shared narratives promotes adaptive decision-making? An initial observation is that the answer must depend on the reason why the shared narrative is propagating – whether it is spreading *because* it facilitated a successful outcome. This could be the case if its adherents are likelier to survive to tell the tale or because the positive outcome motivated them to share the underlying narrative. In contrast, narratives can act as "cultural attractors" without bearing much resemblance to reality, as when they are used to signal group affiliation. **Sheskin et al.** mention research on science communication (e.g., Kahan, Peters, Dawson, & Slovic, 2017), which shows that narrative-confirming beliefs about issues such as climate change and gun control contribute to polarization. Similarly, Caplan (2007) argues that people are often "rationally irrational" in their economic views, which may be wrong from a scientific perspective but useful for gaining social credit with one's in-group, all the while leading to poor government policy through the ballot box.

We view this as a crucial question for future research, and one where formal modeling along the lines suggested by **Jones and Hilde-Jones** may be useful.

## R9. Future directions

The commentators generously provided many ideas for future research. Some commentators provide ideas about how to unify CNT with other approaches such as the Theory of Narrative Thought (**Beach & Wise**), Free Energy Theory (**Friston**), situation models (**Dominey**), and semiotics (**Pelletier et al.**). **Marewski** observes that both cognitive and environmental processes guide decision-making; we agree and would argue that one of CNT's principal insights is that two features of the decision environment are importantly linked: Radical uncertainty and social embeddedness. There are ontogenetic questions about how narratives are used through development (**Tuominen**); we would add the phylogenetic question of how narrative cognition evolves. Interesting questions were raised about the similarities and differences between narratives and other mental representations, such as scripts (**Siegel**), story grammars (**Andreetta et al.**), and situation models (**Dominey**). More work needs to be done to fill in, computationally model, and experimentally test our sketch of narrative representation; suggestions such as those provided by **Caldwell** will be useful, as will incorporating research on lower-level representations as **Newell and Szollosi** suggest. Much is still to be learned about how narratives are used to simulate futures, particularly the question of *which* futures are simulated and how this interacts with the processes described by **Enz and Tamir**. An especially important issue for understanding decision-making across levels of analysis is to model the feedback loops between individual and social processes that contribute to narrative evolution (**Jones & Hilde-Jones**).

Space precludes us from speculating about these fascinating issues in detail; but crucially, it would *be* speculation. We hope that other researchers will continue to take up these important questions so that we need no longer speculate.

In the target article, we closed with our hope that CNT can provide a common vocabulary and motivate a shared set of questions for the decision sciences. The commentaries have fueled our belief that CNT will prove to be, indeed, a useful narrative.

## References


Anderson, J. R. (1990). *The adaptive character of thought*. Erlbaum.

Becker, G. S. (1968). Crime and punishment: An economic approach. *Journal of Political Economy*, *76*, 169–217.

Beckert, J., & Bronk, R. (Eds.) (2018). *Uncertain futures: Imaginaries, narratives, and calculation in the economy*. Oxford University Press.

Caplan, B. (2007). *The myth of the rational voter: Why democracies choose bad policies*. Princeton University Press.

Chater, N. (2018). *The mind is flat: The remarkable shallowness of the improvising brain*. Yale University Press.

Chater, N., & Oaksford, M. (1999). Ten years of the rational analysis of cognition. *Trends in Cognitive Sciences*, *3*, 57–65.

Cushman, F. (2020). Rationalization is rational. *Behavioral and Brain Sciences*, *43*, e28.

Dawkins, R. (1976). *The selfish gene*. Oxford University Press.

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences*, *8*, 307–314.

Fenker, D. B., Waldmann, M. R., & Holyoak, K. J. (2005). Accessing causal relations in semantic memory. *Memory & Cognition*, *33*, 1036–1046.

Fenton O'Creevy, M., & Tuckett, D. (2022). Selecting futures: The role of conviction, narratives, ambivalence, and constructive doubt. *Futures & Foresight Science*, *4*. https://onlinelibrary.wiley.com/doi/full/10.1002/ffo2.111.

Fodor, J. A., & Pylyshyn, Z. (1981). How direct is visual perception? Some reflections on Gibson's "ecological approach". *Cognition*, *9*, 139–196.

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, *11*, 127–138.

Griffiths, T. L., Chater, N., Norris, D., & Pouget, A. (2012). How the Bayesians got their beliefs (and what those beliefs actually are): Comment on Bowers and Davis (2012). *Psychological Bulletin*, *138*, 415–422.

Hirst, W., Yamashiro, J. K., & Coman, A. (2018). Collective memory from a psychological perspective. *Trends in Cognitive Sciences*, *22*, 438–451.

Holyoak, K. J. (1985). The pragmatics of analogical transfer. *Psychology of Learning and Motivation*, *19*, 59–87.

Horne, Z., Muradoglu, M., & Cimpian, A. (2019). Explanation as a cognitive process. *Trends in Cognitive Sciences*, *23*, 187–199.

Johnson, S. G. B., & Ahn, W. (2015). Causal networks or causal islands? The representation of mechanisms and the transitivity of causal judgment. *Cognitive Science*, *39*, 1468–1503.

Johnson, S. G. B., & Ahn, W. (2017). Causal mechanisms. In M. R. Waldmann (Ed.). *Oxford handbook of causal reasoning* (pp. 127–146). Oxford University Press.

Johnson, S. G. B., Jin, A., & Keil, F. C. (2014). Simplicity and goodness-of-fit in explanation: The case of intuitive curve-fitting. In P. Bello, M. Guarini, M. McShane & B. Scassellati (Eds.), *Proceedings of the 36th annual conference of the Cognitive Science Society* (pp. 701–706). Cognitive Science Society.

Johnson, S. G. B., & Keil, F. C. (2014). Causal inference and the hierarchical structure of experience. *Journal of Experimental Psychology: General*, *143*, 2223–2241.

Johnson, S. G. B., Merchant, T., & Keil, F. C. (2020). Belief digitization: Do we treat uncertainty as probabilities or as bits? *Journal of Experimental Psychology: General*, *149*, 1417–1434.


https://doi.org/10.1017/S0140525X22001157 Published online by Cambridge University Press

Johnson, S. G. B., Murphy, G. L., Rodrigues, M., & Keil, F. C. (2019a). Predictions from uncertain moral character. In A. K. Goel, C. M. Seifert & C. Freksa (Eds.), *Proceedings of the 41st annual conference of the Cognitive Science Society* (pp. 506–512). Cognitive Science Society.

Johnson, S. G. B., Rajeev-Kumar, G., & Keil, F. C. (2016). Sense-making under ignorance. *Cognitive Psychology*, *89*, 39–70.

Johnson, S. G. B., Valenti, J. J., & Keil, F. C. (2019b). Simplicity and complexity preferences in causal explanation: An opponent heuristic account. *Cognitive Psychology*, *113*, 101222.

Kahan, D. M., Peters, E., Dawson, E. C., & Slovic, P. (2017). Motivated numeracy and enlightened self-government. *Behavioural Public Policy*, *1*, 54–86.

Keynes, J. M. (1937). The general theory of employment. *Quarterly Journal of Economics*, *51*, 209–223.

Khemlani, S. S., Sussman, A. B., & Oppenheimer, D. M. (2011). Harry Potter and the sorcerer's scope: Latent scope biases in explanatory reasoning. *Memory and Cognition*, *39*, 527–535.

Korman, J., & Khemlani, S. (2020). Explanatory completeness. *Acta Psychologica*, *209*, 103139.

Lagnado, D. A., Waldmann, M. R., Hagmayer, Y., & Sloman, S. A. (2007). Beyond covariation: Cues to causal structure. In A. Gopnik & L. Schulz (Eds.), *Causal learning: Psychology, philosophy, and computation* (pp. 154–172). Oxford University Press.

Lieder, F., & Griffiths, T. L. (2020). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences*, *43*, e1.

Lombrozo, T. (2007). Simplicity and probability in causal explanation. *Cognitive Psychology*, *55*, 232–257.

Lombrozo, T. (2016). Explanatory preferences shape learning and inference. *Trends in Cognitive Sciences*, *20*, 748–759.

Marr, D. (1982). *Vision*. Freeman.

Mills, C. W. (1940). Situated actions and vocabularies of motive. *American Sociological Review*, *5*, 904–913.

Murphy, G. L., & Ross, B. H. (1994). Predictions from uncertain categorizations. *Cognitive Psychology*, *27*, 148–193.

Newell, A. (1973). You can't play 20 questions with nature and win. In W. G. Chase (Ed.), *Visual information processing* (pp. 287–307). Academic Press.

Rottman, B. M., & Keil, F. C. (2012). Causal structure learning over time: Observations and interventions. *Cognitive Psychology*, *64*, 93–125.

Sanborn, A. N., & Chater, N. (2016). Bayesian brains without probabilities. *Trends in Cognitive Sciences*, *20*, 883–893.

Savage, L. J. (1954). *The foundations of statistics*. Wiley.

Scherer, K. R. (2005). What are emotions? And how can they be measured? *Social Science Information*, *44*, 695–729.

Shiller, R. J. (2019). *Narrative economics: How stories go viral and drive major economic events*. Princeton University Press.

Simon, H. A. (1990). Invariants of human behavior. *Annual Review of Psychology*, *41*, 1–19.

Taleb, N. N. (2007). *The black swan*. Random House.

Walasek, L., & Brown, G. D. A. (in press). Incomparability and incommensurability in choice: No common currency of value? *Perspectives on Psychological Science*. https://psyarxiv.com/suw47/.

Woodward, J. (2003). *Making things happen: A theory of causal explanation*. Oxford University Press.

Yin, P., & Sun, J. (2021). Is causation deterministic or probabilistic? A critique of Frosch and Johnson-Laird (2011). *Journal of Cognitive Psychology*, *33*, 899–918.

Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin*, *127*, 3–21.

Zhu, J., & Murphy, G. L. (2013). The effect of emotionally charged information on category-based induction. *PLoS ONE*, *8*, e54286.