Cognitive Attractors and the Cultural Evolution of Religion

Victor Møller Pouslen (vpoulsen@andrew.cmu.edu)
Department of Social and Decision Sciences, Carnegie Mellon University, Pittsburgh, PA 15213 USA

Simon DeDeo (sdedeo@andrew.cmu.edu)
Department of Social and Decision Sciences, Carnegie Mellon University, Pittsburgh, PA 15213 USA
Santa Fe Institute, Santa Fe, NM 87501 USA

Abstract

How do religions evolve? New tools enable the construction of *cultural fitness landscapes*, which provide new perspectives on core questions in the cognitive science of religion. Landscape models, as we show, corroborate longstanding findings, including a strong preference for gods to follow cognitively plausible patterns of agentic behavior, and the idea that the supernatural monitoring by "big gods" can sometimes be assumed by a secular state. They also predict the existence of cultural attractors: under free evolution, a diverse set of religious practices will be channelled into a small number of stable configurations. We observe both "monastic", "evangelical", "indigenous", and "pre-Axial" attractors. Finally, we show the existence of particularly fragile traits: aspects of a religion whose change must either be (1) prevented, through teaching (such as "belief in reincarnation in this world"), or (2) accepted by practitioners as the cost of evolution (such as a loss of state support).

Keywords: religion; cognition; cultural evolution; cultural attractors; machine learning

Introduction

Religion—just as much as any cultural practice—is the art of the possible. Thousands of religions exist, with estimates suggesting that several are founded each day (Norenzayan, 2013). However, the relatively few religions that achieve cultural success, and survive long enough to be broadly recognized, must answer to the cognitive abilities, and predilections, of the individual: what it is possible, or "natural", to believe (Sperber, 1975; Barrett, 2000; Baumard & Boyer, 2013).

Not all aspects of a religion, of course, are reducible to individual cognition. Complex practices, such as monument-building, require joint intentionality and solutions to social dilemmas. The possibility of a religion is, in other words, an emergent property: one that answers to both the logic of an individual's religious experience (James, 1929), and to the wider contexts that enable groups to function together (Norenzayan et al., 2016).

Making sense of how these cognitive and ecological constraints interact is not a trivial matter. One way to approach the question is to focus on one or two aspects of a practice, and to look for more fundamental models of how they function either for the individual agent, or in a group context. Luhrmann et al. (2021), for example, showed that the direct experience of spiritual entities—a key aspect of many religions—emerges from an individual's adoption of particular folk theories of other minds, along with a predilection for absorption into sensory experience. Along more social dimensions, Power (2017), for example, showed that costly religious rituals can help establish an individual's pro-social reputation, making them potentially adaptive in societies that require complex, but *ad hoc*, cooperation.

This approach contrasts with the data-driven approach to cultural evolution that we follow here, which relies on large-scale datasets where religions are coded on a number of distinct aspects. In this case, one uses the data to infer a *cultural landscape* that captures how aspects of a religion (*e.g.*, whether or not it demands participation in large-scale rituals) interact with each other to make the final configuration more or less probable (what we might think of as "stable", "coherent", "natural", or "attractive").

Given a sufficiently diverse set of religions and questions, a cultural landscape model can capture both endogenous (e.g., cognitive) and exogenous (e.g., socio-political) constraints on religious practices. Both endogenous, and exogenous, features have been emphasized in the literature on the cultural evolution of religion. Barrett (2000)'s classic work, for example, emphasized that religious concepts tend to be minimally counterintuitive relative to a person's overall conceptual scheme. Under this account, the supernatural agents in a religion ought to "look like" more ordinary agents. *Which* agents they resemble, of course—tricksters or norm enforcers, vengeful children or wise elders—might vary a great deal, but we expect the patterns of a theology to track, all other things being equal, patterns in the mundane. A parallel tradition, beginning with Weber (1904), emphasises the extent to which these cognitive features co-evolve within a wider social context.

As we shall see, cultural landscapes also make predictions about the dynamics of religious practice; we can approach both (1) how a practice might be expected to transform over time, under the influence of cultural attractors (Miton, 2022), and (2) "points of fragility"—the aspects of a religion that are the least constrained by other features and thus the most likely to be either explicitly taught (Miton & DeDeo, 2022), or the most dependent on contingent social factors.

The Idea of a Cultural Landscape

The landscape approach we adopt here begins with a set of N properties of interest; for the sake of simplicity, and to match the predominant format of the *Database of Religious History*

(DRH; Slingerland & Sullivan, 2017; Slingerland, Monroe, & Muthukrishna, 2022), we take these properties to be answers to binary YES/NO questions. In our curated subset of the DRH, for example, one question is whether religions have a belief in "reincarnation in this world"; the full list of questions we use is at the end of the paper. A complete set of YES/NO answers to the question set is called a configuration.

A cultural landscape is built from this data: it is a parsimonious model of the underlying distribution that (in the long timescale limit) the historical record is sampling from. Because of the heterogeneity in the types of questions that are included in the present analysis, the distribution captures both cognitive, and social, constraints that make some configurations more or less likely than others, independent of any particular genealogy. This inference step will, in principle, provide us with a slice of the theoretically "true" landscape—the idealized, complete account that would take into account all of the causally relevant features. However, as the number of features increase, the space of configurations expands exponentially, and in practical work, we are forced to work with lower dimensional projections, a partial reconstruction built from a curated sub-selection of the potential features that constitute a (religious) cultural practice.

Even a partial reconstruction is radically undersampled by religions in the historical record, and inferring the full distribution of the cultural landscape is a non-trivial challenge. We use a "maximum entropy" model, an approach that has seen recent use in both the cognitive, social, and biological sciences (Lee, Broedersz, & Bialek, 2015; Louie, Kaczorowski, Barton, Chakraborty, & McKay, 2018; Stephens, Osborne, & Bialek, 2011), where the probability of a configuration is the emergent consequence of complex combinations of pairwise constraints; formally, for a configuration $\{\sigma_1,\ldots,\sigma_N\}$, where a YES (NO) to question i is represented by σ_i equal to $+1$ (-1, respectively). This "global" probability is given by

$$P(\{\sigma_1,\ldots,\sigma_N\}) = \frac{\exp\left(\sum_{ij} J_{ij}\sigma_i\sigma_j + \sum_i h_i\sigma_i\right)}{Z}, \qquad (1)$$

where J_{ij} are the pairwise constraints, h_i, the "local fields", help fix the average values for each property, and Z is a normalization constant. Despite its simplicity, Eq. 1 can capture complex patterns and higher-order effects (Schneidman, Berry, Segev, & Bialek, 2006); it is sometimes known as a "Boltzmann machine" (Ackley, Hinton, & Sejnowski, 1985) or a "Hopfield network" (Hillar, Sohl-Dickstein, & Koepsell, 2012).

Cognition in an ecological context

Once the J_{ij} and h_i are inferred from data, Eq. 1 becomes a powerful tool for investigation. Consider, for example, two core features of the "big gods" debate (Norenzayan, 2013; Norenzayan et al., 2016): that god is watching (monitoring), and that god is punishing (which often coincides with gods that have a general moral concern, and also reward compliance (Norenzayan et al., 2016)). Recent work by Poulsen and DeDeo (2023) provides a landscape model over these two features, along with eighteen additional ones, inferred from 407 religious groups in a subsample of the DRH. The algorithm (Poulsen & DeDeo, 2023) allows for the integration of partial data, providing predictions for these missing values after estimation; in the work presented here, when we refer to the configuration associated with an observed group that has missing data, we use the maximum likelihood estimate from the inferred landscape model.

Given the inferred parameters, Eq. 1 can be used to estimate how the two "big gods" features co-vary; holding everything else equal, and averaging over all 2^{20} possible religious configurations, we have

(Full Landscape)	Punishing	Not Punishing
Monitoring God	0.427	0.089
Not Monitoring God	0.083	0.400

i.e., the landscape model shows how monitoring and punishing tend to be coupled—as might be expected if god is understood through a cultural model of human agency (Barrett, 2000), where human agents are expected to not only watch, but also have goals and desires. Abstracting away from biases induced by historical evolution, we find that "big gods" and non-"big gods" religions are roughly equally preferred. Meanwhile, the "off-diagonal" cases—"otiose" gods, who watch but do not punish, and "non-surveilling" gods, who do not watch, but do punish—are a factor of five times less likely. The discussion in Barrett (2000) suggests that these cases might violate intuitions that we have from analogizing gods to more mundane, human agents: a powerful human agent, for example, that care enough about norm violations to punish them, would also be expected to monitor whether transgressions occurred. Following a similar line of reasoning, any agent that is monitoring, would be expected to have intentionality, which might take the form of meting out punishment. The landscape model provides new support for this claim.

The landscape model also enables us to examine the contextual effects that moderate the sense-making drives that regularize god-like behavior to familiar agentic patterns. In particular, we can go through the remaining 18 questions, and see which one produces the largest effect on the probability distribution: formally, if $\vec{\sigma}$ represents the pair of "big gods" questions, we look for the question i with the largest Jensen-Shannon distance between $P(\vec{\sigma},\sigma_i=+1)$ and $P(\vec{\sigma},\sigma_i=-1)$. This turns out to be whether or not the religious group enjoys state support. When we restrict to the "state-supported" area of the landscape, we get

(state-support YES)	Punishing	Not Punishing
Monitoring God	0.404	0.052
Not Monitoring God	0.122	0.433

In particular, our landscape model suggests that cognitively counterintuitive gods who punish but do not monitor become more preferred when accompanied by official state support:

informally, state support appears to serve as a substitute for surveillance theology. In the ancient Ugarit religion (Portis, 2023), for example, norm violations may be brought before god as one might bring a case before a state official for trial—*i.e.*, the surveillance is carried out by society, while the god dispenses justice. Conversely, the landscape suggests that state support makes "otiose" gods (those who watch, but do not punish) less likely. The possibility of such substitutions (god takes on part of the role of state; state performs part of the role of god) corroborates Norenzayan (2013); Norenzayan et al. (2016) where it is argued that big gods might have initially enabled large-scale societal cooperation, but that secular institutions (e.g., courts, police) can assume the role of enabling prosocial behavior in contemporary societies.

Fitness versus (Meta)stability

The previous section considered a three-way interaction, but Eq. 1 learned on the DRH data provides a distribution over a 20-dimensional space. If we think of the probability of each configuration as having a "height", we can draw on metaphors from geographical landscapes in physical space to approach an intuition about the landscape topography. For example, configurations that are unlikely are said to be in low-lying "valleys", while more probable configurations might be though of as "hills" or "peaks".

These ideas are a central feature in evolutionary biology, which has a long tradition of talking in terms of fitness landscapes (Pitzer & Affenzeller, 2012; Fragata, Blanckaert, Louro, Liberles, & Bank, 2019). A key insight from this line of work is that while the global probability of a configuration (*i.e.*, Eq. 1) might measure overall (sociocognitive) fitness on very long timescales, it does not paint a complete picture of the landscape. To understand the evolution of a religion over time, we need to broaden this conception, because the presence of other probable "nearby" configurations—configurations that differ in only one or a few attributes—can make even a high-probability configuration less stable, because the neighbours serve both as an attractive, and a cognitively and socially accessible, alternative. Conversely, a less probable point in the landscape that is surrounded by even less probable neighbours may achieve stability: it may not satisfy the cognitive and social constraints very well, but it will benefit from a lack of nearby alternatives that satisfy the constraints better.

An explicit model for this effect is provided by Glauber dynamics (Glauber, 1963); given a particular base configuration i, if we restrict exploration to a set of neighbours, $\mathcal{N}(j)$, the probability that we move from i to one of those neighbours j is given by

$$P(i \to j) = \frac{1}{|\mathcal{N}|} \frac{P(j)}{P(i)+P(j)}. \qquad (2)$$

A natural choice for the neighbour set $\mathcal{N}(j)$ is the configurations that differ in only one question; this corresponds to the idea that religions change gradually, over time—either in response to social changes (*e.g.*, the loss of official political support) or more basic cognitive shifts (*e.g.*, the adoption of the idea that a god might practice surveillance).

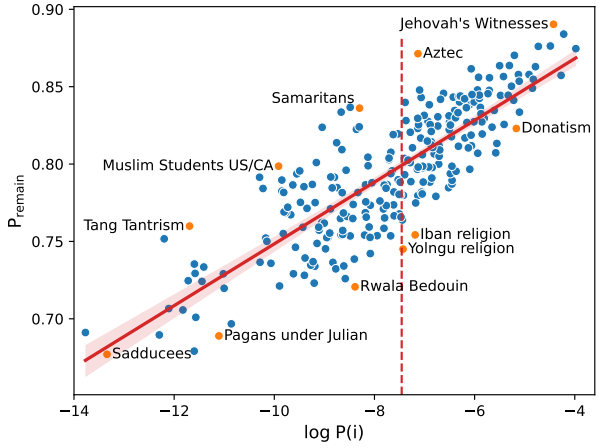

Figure 1: Fitness is not (necessarily) stability. Some religions (*e.g.*, Donatism) are relatively high in overall fitness, but have other attractive neighbours. Others (*e.g.*, the Samaritans) may have lower overall fitness, but are located in regions of the landscape where nearby alternatives are significantly less attractive. Vertical (dotted) red line is the median $\log P(i)$ for the 260 maximum likelihood configurations. Solid red line shows the best fit and 95% confidence intervals.

We can then compare the global probability of a configuration, $P(i)$, to its local stability, $1-\sum_j P(i \to j)$, or P_{remain}, for short. This is shown, for the observed religions in our DRH set, in Figure 1 (because some religions have the same configuration, the 407 observed religions map to 260 distinct configurations). While there is a clear connection between high fitness (being globally preferred), and local stability, there is also a substantial spread around the trend line. Configurations that fall below the line are less stable than expected given their probability, while configurations above the line are more stable than expected.

Figure 1 natually divides observed religious practices into four groups. A religion may be found at an "isolated peak": high probability, and surrounded by configurations that tend to be less probable. These are found in the top right corner, with the "Jehovah's Witnesses", and the "Aztec Imperial Core" being two examples. Other high-fitness religions are part of a "mountain range"; high probability, but surrounded by other probable configurations. These include "Donatism", "Iban traditional religion", and "Yolngu religion".

Among the less favored configurations, we observe both "valleys", and "hillocks". A configuration with low probability, which is additionally surrounded by relatively more probable configurations, is in a valley, and may tend to flow upwards in the landscape; examples of these cases (bottom left quadrant) include the "Sadducees", "Pagans under Julian", and the "Rwala Bedouin". Finally, there are hillocks—low-

lying locally preferred configurations such as the "Samaritans", "Muslim Students in the US and Canada", and "Tang Tantrism". These configurations are sometimes referred to as metastable (Cortês, Kauffman, Liddle, & Smolin, 2022): on short timescales, they are expected to be resistant to changes, but—given their lower overall probability—may, eventually, evolve by longer, more unlikely, macroevolutionary leaps.

Cultural Attractors

The analysis of the previous section reveals a complex relationship between local, and global, stability. As this dynamical process unfolds, we expect religious groups to evolve towards a smaller set of cultural attractors, local maxima in the landscape, where further evolution is disfavored. To study this in greater detail, we use Eq. 2 as a simple, Markovian, evolution rule that is consistent, on very long timescales, with the global fitness landscape, and simultaneously unstructured enough to capture the variety of mechanisms, such as regularization (Ferdinand, Kirby, & Smith, 2019) and transformation (Miton, 2022), that we expect to underlie this drive in real-world situations.

In the high-dimensional spaces that characterize cultural landscapes, forward evolution can lead to multiple attractors. Figure 2) shows one example, for the case of (contemporary) Irish Catholicism. In order to understand the shorter-timescale dynamics, we show only the paths leading to higher-likelihood configurations, and do not consider paths through configurations of lower probability which, under Eq. 2, are possible but (usually strongly) dispreferred. Flow is visualized as going "upwards" on the page, with the vertical position tracking relative log-probability. Beginning with a focal configuration, we show the most likely uphill transitions over the landscape. When a configuration corresponds to at least one observed religion, we choose one of those religions as its label; in the Irish Catholicism case, the forward evolution also involves passage through configurations that (despite being higher likelihood than Irish Catholicism) have no matches in our observed data; these are colored light blue.

The path from Irish Catholicism eventually terminates at one of three local maxima, colored in red. In descending order of probability, the terminal configurations are associated with (1) the Medieval Cistercian Order (and a number of other monastic orders, Christian and non-Christian); (2) the Jehovah's Witnesses (and a number of evangelical Protestant religions, such as the Anabaptists and Churches of Christ); and (3) the Sokoto Caliphate, a religious theocracy. The heavy black edges trace the maximum likelihood path: the most likely transition from the Irish Catholics is to the Yiguan Dao configuration, and from there to the Cistercian configuration. This path involves two bit-flips: first, the group must acquire the "supernatural punishment" trait, then the "special corpse treatment" trait.

These landscapes are complex, and the higher-order interactions between traits lead to interesting path-dependent effects. Another path to the Cistercian configuration goes via a configuration with no corresponding religion in our data, to the Qumran Movement, and through either Opus Dei or the Yiguan Dao, before again terminating at the Cistercian configuration. In both of these cases, the path, which always moves uphill, involves both the loss of the "official political support" trait, and its later reacquisition. This shows that we should not, in most cases, think of traits as being globally favored or disfavored: cultural evolution is not a slow acquisiton of unambiguously more useful traits, but a context-sensitive navigation over a landscape of competing social and cognitive demands.

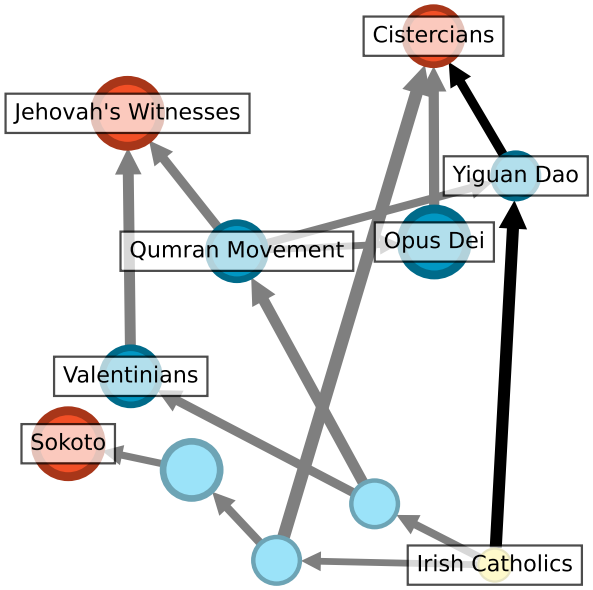

Figure 2: The evolutionary paths from "Irish Catholicism". Vertical position based on $log\ P_i$. Light blue nodes are configurations that do not correspond to an observed religion, dark blue nodes correspond to at least one observed religion, and red nodes are configurations that are more probable than any of their immediate neighbors. Node size is scaled by the Hamming distance (plus a constant) from the Irish Catholic configuration. Edge size is scaled by the relative probability of transition, and the black edges follow the "naive" highest probability path.

Fig. 2 has three terminal attractor states. In Table 1, we list the eleven attractors found as terminal states for the 407 religious groups (corresponding to 260 configurations) in our dataset. We recognize the Cistercian group, the Jehovah's Witnesses group, and the Sokoto, that we encountered as potential termination states for the Irish Catholics; there are others, including a number of configurations strongly associated with pre-Axial traditions (Bellah, 2017).

Table 1 gives further insight into the path dependent na-

N	$P(i)$	Characteristics	Example
83	0.018716	Monastic	12th c. Cistercians
42	0.007348	Indigenous	Iban religion
29	0.011989	Evangelical	Jehovah's Witnesses
26	0.014685	Pre-Axial	Ancient Egypt
17	0.001779	⋆	Sokoto Caliphate
11	0.002316	⋆	Pythagoreanism
10	0.005486	Evangelical	Free Methodist Church
9	0.005943	Pre-Axial	Ancient Thessalians
8	0.002682	⋆	Messalians
7	0.001617	⋆	Hidatsa
2	0.000800	Pre-Axial	Aztec Imperial Core

Table 1: The eleven cultural attractors inferred from the DRH. N refers to the number of configurations that have this as their maximum-likelihood path attractor; $P(i)$ the global probability (Eq. 1). Many attractor configurations are associated with multiple religions in the data; "Characteristics" provides a rough description; where there is only one observed religion this is noted by ⋆.

ture of cultural evolution. There is reasonable correlation between the global probability of a configuration, $P(i)$, and the number of observed religions that terminate at this configuration, N, but the relationship is not straightforward: in general, microevolutionary steps lead to high-fitness peaks. However, the "actually existing" religions in our sample occupy regions of the space that lead them, for example, to terminate at the "indigenous" group that includes the Iban more often than at either the evangelical group that includes Jehovah's Witnesses, or the "pre-Axial" group that includes Ancient Egyptian traditions; this is despite the fact that the latter two groups better satisfy the landscape constraints.

Two effects enable a less preferred configuration to serve as a more powerful attractor than a more (globally) preferred configuration. First, the peak may have a larger basin of attraction, meaning that there are more configurations such that purely up-hill walks will take you to that peak rather than others. Second, the historical data (and, more generally, the geneaological effects of human cultural evolution) may have led religions to concentrate in one peak's basin of attraction more than others.

Fragile Traits and Religious Stability

The previous result concerns how religious practices might evolve under cognitive and social pressures. In general, religions are expected to change over time, drawn towards a small number of cultural attractors. In this section, we examine the phenomenon of *fragile traits*—features that are most easily varied—and the related idea that religions might achieve stability by fixing (*e.g.*, through explicit teaching) particular features to their current value, an idea previously explored in theoretical work (Miton & DeDeo, 2022).

A simple way to examine this is to look at how local stability (P_{remain}) changes, for a particular group, as we fix different features—if we "forbid" the religion to vary in certain traits. A simple way to do this is to first fix the "most fragile" trait—the feature associated with the most likely one-question change under Eq. 2—and see how this shifts P_{remain}. This is shown in Figure 3, where we fix the most fragile trait, then additionally the second most fragile trait, until all but one trait is fixed.

As we fix more and more traits, local stability grows, and the distribution narrows. While the initial spread in stability is large (with no traits fixed, median stability is 80% with 95% HDI intervals between 72% and 88%), this variability shrinks as more traits are fixed (*e.g.*, for 10 fixed traits the median stability is 97% with HDI intervals between 94% and 98%). Even very "unstable" religions, such as Paganism under Julian or Santal (shown in Figure 3), can be stabilized in this exogenous fashion.

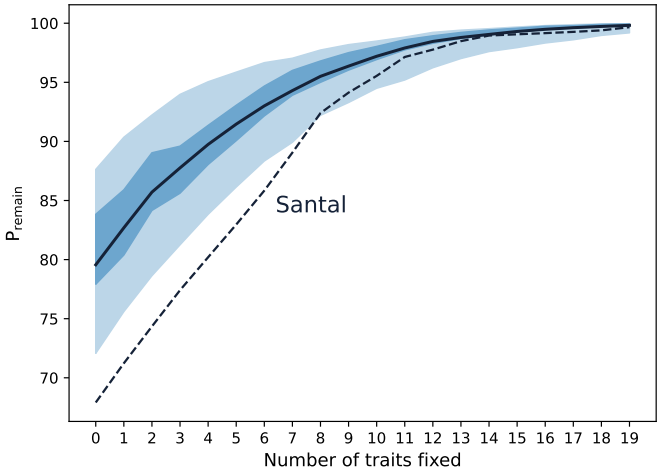

Figure 3: Local stability (P_{remain}) for the 260 observed configurations, varying the number of traits fixed. Central line is the median, while the wide band is 95% highest density interval (HDI), and the narrow band is 50% HDI interval. Religions that appear at unstable points on the landscape (*e.g.*, Santal) can be made stable by interventions on a fraction of the traits.

Contextual effects mean there will be variation, from religion to religion, in which traits are the most fragile. However, our model implies that some features are fragile across many religions. In descending order, the most fragile traits are "reincarnation in this world" (appears in the top five most fragile traits for 54% of the observed religions), "official political support" (53%), "monumental religious architecture" (47%), "grave goods" (41%), and "special treatment for corpses" (40%).

The most fragile trait in the landscape is an epistemic one; our landscape model suggests that there is something particularly difficult about preserving a belief in (or, almost as difficult, a non-belief in) worldly reincarnation. This is a claim with cognitive consequences—our model suggests that, across various religious sets of beliefs and practices, it is rela-

tively easy to shift one's belief in worldly reincarnation compared to, say, the idea that god might punish you for norm violations.

This first result is also, potentially, a prediction about how religions manage the threat of change. If a religion wishes to remain unchanged, then we would expect it to pay attention to this more fragile trait. One way to fix a trait, of course, is through explicit teaching; as mischievously suggested by the atheist Voltaire, the Jesuits said, "give me the child for the first seven years and I will give you the man".

While our model points out a special role for reincarnation, the remaining most-fragile traits are of a rather different nature. Two of them, "grave goods" and "special treatment for corpses", are behavioral practices with (potentially) theological implications. It is perhaps not surprising that, of the behavioral patterns, they are the most easily lost; in both cases, they concern practices that are only infrequently undertaken (burials are, usually, far rarer than daily or weekly rituals).

The two remaining traits are "official political support" and "monumental religious architecture", which concern the interaction between the religion and large-scale social behaviors and systems. Under our formal definition, of course, a religion that changes status on any questions has "changed".

These cases highlight that, although change is usually associated with important shifts in identity and self-conception, this might not always be the case. To take an extreme example, Danish Lutherans technically enjoy state support; but it is unlikely that they would consider themselves distinct in any important fashion from German Lutherans, who do not. In other cases, religions might indeed tie their identity to these more exogenous features—for many pre-Axial religions, for example, the political leader has an intimate relationship with the supernatural almost by definition (Bellah, 2017). The centrality of different practices for different cultures can vary drastically, and in many cases "fragility" might be better conceptualized as "fluidity".

Discussion

Foundational work in the cultural evolution of religion has often taken a "top-down" approach, starting with a hypothesized interaction between a small number of key cognitive or social mechanisms, and validating them through case studies, laboratory experiments, or cross-cultural surveys.

This work takes a different, "bottom-up" and data-driven approach, surveying at once a much larger set of features, inferring their complex interactions (the J_{ij} structure) in a synoptic fashion, and summarizing them in the metaphor of a probabilistic landscape. While some previous work in the cultural attractor literature has integrated large-scale databases as an initial step towards establishing factors of attraction (Miton, Claidière, & Mercier, 2015), the approach taken in this work represents a more ambitious approach, moving beyond an initial assessment of the predominance of traits in the cultural record, and towards a coherent and increasingly complete picture of both the interactions that lie beneath, and their consequences.

By contrast, the most ambitious top-down approaches (*e.g.*, Norenzayan et al. (2016)) are necessarily restricted to lower dimensional projections of this landscape—such approaches yield fruitful and interesting contributions, but are challenged by how cultural systems rely critically on complex relationships, and higher-order correlations, to maintain stability or, conversely, facilitate evolvability—aspects that are lost when the landscape is not inferred together.

A natural objection is that our work is similarly limited. Religious cultures may not be dictated by two or three interactions, but neither are they as simple as the 20-dimensional objects we consider here. This is both a concern, and a challenge: the inference process grows exponentially harder for additional questions, and there is some practical limit to how far we can go. A combined approach, that includes relying on fundamental theories to constrain some relationships *a priori*, based on the best, current, theoretical understanding of relationships, is likely the best way forward. Coarse-graining, or "renormalization", is necessary, but never theory-free: couplings at the coarse-grained scale will often tend to be lower than are at more fundamental levels.

Despite these limitations, the approach here, as we have seen, reveals a great deal. Typically, and perhaps both naturally and appropriately, cultural evolution is considered with a view toward temporal evolutionary trends from an "old" system to a "new" one. One important contribution is to help dispel the notion that religious evolution has a directedness that climbs steadily to a unique solution. There are many attractors, and our results suggest that while certain modern phenomena, such as the emergence of non-state evangelical religions, are natural (temporary) resting points, far older models remain as equally-valid places for future evolution.

For example, configurations with neither monitoring, or punishing, gods, have roughly the same stability (on average) than configurations with these "big gods" traits. This suggests that while "big gods" might become favored in societies where they provide benefits to the group, a culture might later return to configurations without these traits, if other cultural facets (*e.g.*, courts or police) can fill these needs: modern solutions that replace a monitoring god with a monitoring state (such as modern Scandanavia) can be considered, not an innovation, but a return to more ancient paradigms found, for example, in ancient Ugarit.

Question List. Are supernatural beings present? Is supernatural monitoring present? Do supernatural beings mete out punishment? A spirit-body distinction? Belief in afterlife? Reincarnation in this world? Does the religion have official political support? Does the religious group have scriptures? Monumental religious architecture? Special treatments for adherents' corpses? Are co-sacrifices present in tomb/burial? Are grave goods present? Are formal burials present? Does membership in this religious group require castration? ..sacrifice of adults? ...sacrifice of children? ...self-sacrifice (suicide)?

References

Ackley, D. H., Hinton, G. E., & Sejnowski, T. J. (1985). A learning algorithm for boltzmann machines. *Cognitive Science*, *9*(1), 147–169.

Barrett, J. L. (2000). Exploring the natural foundations of religion. *Trends in Cognitive Sciences*, *4*(1), 29-34. doi: https://doi.org/10.1016/S1364-6613(99)01419-9

Baumard, N., & Boyer, P. (2013). Explaining moral religions. *Trends in cognitive sciences*, *17*(6), 272–280.

Bellah, R. (2017). *Religion in human evolution: From the paleolithic to the axial age*. Harvard University Press.

Cortês, M., Kauffman, S. A., Liddle, A. R., & Smolin, L. (2022). *Biocosmology: Biology from a cosmological perspective.* Retrieved from `https://arxiv.org/abs/2204.09379` doi: 10.48550/ARXIV.2204.09379

Ferdinand, V., Kirby, S., & Smith, K. (2019). The cognitive roots of regularization in language. *Cognition*, *184*, 53-68. doi: https://doi.org/10.1016/j.cognition.2018.12.002

Fragata, I., Blanckaert, A., Louro, M. A. D., Liberles, D. A., & Bank, C. (2019). Evolution in the light of fitness landscape theory. *Trends in ecology & evolution*, *34*(1), 69–82.

Glauber, R. J. (1963). Time-dependent statistics of the ising model. *Journal of mathematical physics*, *4*(2), 294–307.

Hillar, C., Sohl-Dickstein, J., & Koepsell, K. (2012). *Efficient and optimal binary hopfield associative memory storage using minimum probability flow.* (Neural Information Processing Systems (NurIPS) workshop on Discrete Optimization in Machine Learning (DISCML)) doi: 10.48550/ARXIV.1204.2916

James, W. (1929). *The varieties of religious experience: A study in human nature*. New York: Cambridge University Press.

Lee, E. D., Broedersz, C. P., & Bialek, W. (2015). Statistical mechanics of the us supreme court. *Journal of Statistical Physics*, *160*(2), 275–301.

Louie, R. H. Y., Kaczorowski, K. J., Barton, J. P., Chakraborty, A. K., & McKay, M. R. (2018). Fitness landscape of the human immunodeficiency virus envelope protein that is targeted by antibodies. *Proceedings of the National Academy of Sciences*, *115*(4), E564-E573. doi: 10.1073/pnas.1717765115

Luhrmann, T. M., Weisman, K., Aulino, F., Brahinsky, J. D., Dulin, J. C., Dzokoto, V. A., ... Smith, R. E. (2021). Sensing the presence of gods and spirits across cultures and faiths. *Proceedings of the National Academy of Sciences*, *118*(5), e2016649118. doi: 10.1073/pnas.2016649118

Miton, H. (2022). Cultural attraction. *PsyArXiv*. (To appear in *Oxford Handbook of Cultural Evolution*.)

Miton, H., Claidière, N., & Mercier, H. (2015). Universal cognitive mechanisms explain the cultural success of bloodletting. *Evolution and Human Behavior*, *36*(4), 303–312.

Miton, H., & DeDeo, S. (2022). The cultural transmission of tacit knowledge. *Journal of the Royal Society Interface*, *19*(195), 20220238.

Norenzayan, A. (2013). *Big gods: How religion transformed cooperation and conflict*. Princeton University Press.

Norenzayan, A., Shariff, A. F., Gervais, W. M., Willard, A. K., McNamara, R. A., Slingerland, E., & Henrich, J. (2016). The cultural evolution of prosocial religions. *Behavioral and Brain Sciences*, *39*, e1.

Pitzer, E., & Affenzeller, M. (2012). A comprehensive survey on fitness landscape analysis. *Recent advances in intelligent engineering systems*, 161–191.

Portis, R. (2023, Feb). *Ugarit.* Retrieved from `https://religiondatabase.org/browse/938/`

Poulsen, V. M., & DeDeo, S. (2023). Inferring cultural landscapes with the inverse ising model. *Entropy*, *25*(2). doi: 10.3390/e25020264

Power, E. A. (2017). Discerning devotion: Testing the signaling theory of religion. *Evolution and Human Behavior*, *38*(1), 82-91. doi: https://doi.org/10.1016/j.evolhumbehav.2016.07.003

Schneidman, E., Berry, M. J., Segev, R., & Bialek, W. (2006). Weak pairwise correlations imply strongly correlated network states in a neural population. *Nature*, *440*(7087), 1007–1012.

Slingerland, E., Monroe, M. W., & Muthukrishna, M. (2022). The database of religious history (drh): Ontology, coding strategies and the future of cultural evolutionary analyses. *Religion, Brain and Behavior*. (Submitted)

Slingerland, E., & Sullivan, B. (2017). Durkheim with data: The database of religious history. *Journal of the American Academy of Religion*, *85*(2), 312–347.

Sperber, D. (1975). *Rethinking symbolism*. Cambridge University Press. (Translated by Alice L. Morton.)

Stephens, G. J., Osborne, L. C., & Bialek, W. (2011). Searching for simplicity in the analysis of neurons and behavior. *Proceedings of the National Academy of Sciences*, *108*(supplement_3), 15565–15571.

Weber, M. (1904). *The protestant ethic and the spirit of capitalism*. Dover Publications. (Contemporary edition, 2012.)