Attentional resource allocation and cultural modulation in a computational model of ritualized behavior

Kristoffer L. Nielbo\textsuperscript{a,b,c} and Jesper Sørensen\textsuperscript{a,b,c}

\textsuperscript{a}Department of Culture and Society, Aarhus University, Aarhus, Denmark; \textsuperscript{b}Religion, Cognition and Culture, Aarhus University, Aarhus, Denmark; \textsuperscript{c}Interacting Minds Centre, Aarhus University, Aarhus, Denmark

\textbf{ABSTRACT}

How do cultural and religious rituals influence human perception and cognition, and what separates the highly patterned behaviors of communal ceremonies from perceptually similar precautionary and compulsive behaviors? These are some of the questions that recent theoretical models and empirical studies have tried to answer by focusing on ritualized behavior instead of ritual. Ritualized behavior (i.e., a set of behavioral features embedded in rituals) increases attention to detail and induces cognitive resource depletion, which together support distinct modes of action categorization. While ritualized behaviors are perceptually similar across a range of behavioral domains, symbolically mediated experience-dependent information (so-called cultural priors) modulate perception such that communal ceremonies appear coherent and culturally meaningful, while compulsive behaviors remain incoherent and, in some cases, pathological.

In this study, we extend a qualitative model of human action perception and understanding to include ritualized behavior. Based on previous experimental and computational studies, the model was simulated using instrumental and ritualized representations of realistic motor patterns and the simulation data were subjected to linear and non-linear analysis. The results are used to exemplify how action perception of ritualized behavior (a) might influence allocation of attentional resources and (b) can be modulated by cultural priors. Further explorations of the model show why behavioral experiments might fail to capture modulation effects of cultural priors and that cultural priors in general reduce the chaoticity of time-dependent action processing.

\textbf{KEYWORDS}

Action understanding; expectation modulation; ritualized behavior; rituals; simulation

\section{1. Introduction}

The scientific study of rituals is all too often hampered by a failure to reach an agreed-upon definition or, more moderately, an empirical stipulation of what exact phenomena should fall under the concept. Several contemporary researchers, therefore, propose to focus on ritualized or non-functional behaviors, which in comparison to “rituals” are narrowly defined theoretical constructs and more accessible to qualitative and quantitative modeling (Boyer & Liénard, 2006; Eilam, Zor, Szechtman, & Hermesh, 2006; Nielbo & Sørensen, 2011; Zor, Hermesh, Szechtman, & Eilam, 2009). At the level of behavior, these constructs (henceforth “ritualized behavior”) target a specific set of behavioral features that are embedded in individual and collective rituals. This set of behavioral features consists of...
rigidity, scriptedness, stereotypy, redundancy, and goal-demotion (Boyer & Liénard, 2006), which together generate behavior that is perceived as causally intransparent (Sørensen, 2007).

It is a common characteristic of the features of ritualized behavior that they elicit an increased attention to detail and correct performance (Boyer & Liénard, 2006) and they therefore produce specific perceptual and cognitive effects. Since attentional resources are limited, tasks that utilize attentional resources are in a constant resource competition (Engle, 2002; Engle, Conway, Tuholski, & Shisler, 1995). Attention to detail and correct performance is computationally demanding and, therefore, depletes attentional resources, making them temporarily unavailable for other tasks. Performance of ritualized behavior, therefore, causes cognitive resource depletion (Boyer & Liénard, 2006; Schjoedt et al., 2013). In anthropology, the scope of attention associated with ritual performances is sometimes described as local and referred to by means of a tunnel analogy (Bloch, 1989). According to the tunnel analogy, performance of a ritual is comparable to passing through a narrow cylindrical shape that prevents global assessments of direction, duration, and size. Ritualized behavior embedded in ritual performances is a likely explanatory candidate for a local scope of attention. At the level of categorization, attention to detail and resource depletion resulting from embedded ritualized behavior might also contribute to an explanation of the widespread distinction between ritual and instrumental behavior. If ritualized behavior depletes perceptual and cognitive resources, the behavior is likely to be experienced as perceptually salient in comparison to frequent instrumental behaviors and, therefore, to be linguistically encoded as a distinct class (Nielbo & Sørensen, 2013). A difference in resource allocation and associated experience will, all things being equal, generate different modes of memory encoding and integration (Whitehouse, 2002; Xygalatas et al., 2013).

Many daily practices and ceremonies inside as well as outside religious contexts involve ritualized behavior, and several of the highly repetitive behaviors exhibited by children, pregnant women, and patients suffering from obsessive-compulsive disorder (OCD) and obsessive-compulsive spectrum disorders are also of a ritualized character (Boyer & Liénard, 2006; Lienard & Boyer, 2006). That ritualized behavior belongs to very different behavioral domains – for instance, child-caring, religious and pathological domains – points to the role cultural expectations or priors1 have in modulating human perception and cognition. Due to the overt similarity of ritualized behaviors, the ritual lighting of a sacrificial fire by an Adhvaryu priest in the Vedic sacrifice might to the outsider seem similar to the compulsive checking and ordering by an OCD patient (Dulaney & Fiske, 1994; Fiske & Haslam, 1997). To the proficient cultural agent, however, there is an abysmal difference between the two behaviors. The sacrificial ritual is socially sanctioned, culturally meaningful, and can serve several positive signaling functions, while the compulsive behavior is pathological, culturally incoherent, and loaded with negative valence. Because the ritualized aspects of the two behaviors are overtly similar, their difference in categorization is likely to be found in the cultural priors – that is, symbolically mediated experience-dependent information – that are associated with the behaviors. Although both agents process qualitatively similar sensory information, the percept of the culturally proficient agent is modulated by a cultural prior that underlies [the formation of] action expectations. The outsider lacks this cultural prior and must rely on the available sensory information only. Prior-based expectation modulation of sensory information should, in other words, be one of the primary mechanisms through which ritualized behavior becomes culturally meaningful.

Taken as a whole, ritualized behavior is a central component of both communal and individual rituals. It has distinct perceptual and cognitive effects that support a conceptual distinction between ritual and instrumental behavior; that is, ritualized behavior facilitates categorization of behaviors as exactly rituals. Further, within the category of ritual, expectation modulation through cultural priors plays an important role in selecting ritualized behaviors that although similar in terms of behavioral features are different in terms of meaning and value. With this study we wanted to qualitatively model perceptual and cognitive effects of ritualized behavior as well as cultural priors in action processing and categorization.
1.1. Ordinary action perception

Because ritualized behavior is a type of action, a model of ritualized behavior assumes a theory of human action processing (McCauley & Lawson, 2002). Humans are equipped with an action parsing system that is responsible for action perception and understanding (Zacks & Sargent, 2010; Zacks, Speer, Swallow, Braver, & Reynolds, 2007; Zacks & Tversky, 2001). This system parses continuous streams of perceptual information into discrete action units by operating at distinct levels. At a coarse level, parsing consists of placing boundaries between actions, thereby differentiating one action from another by linking motor components into goal-oriented sequences (e.g., making coffee). At a fine level, the parsing system differentiates one motor component from another by placing boundaries within an action sequence, which specifies the causal sub-components of an action (e.g., lifting the kettle/turning on the water/filling the kettle …). The action parsing system is constantly engaged in predicting an action’s time-dependent structure. Action parsing – that is, the placement of action boundaries – is a function of this prediction process. Somewhat simplified, the system utilizes three computational resources to predict an action’s development, namely sensory input, priors, and monitoring of prediction error. If the priors – that is, stable expectations of the action’s structure – correspond to the sensory input, the system is accurate in predicting the time-dependent structure of the action, there are little or no prediction errors (e.g., pouring coffee into the coffee-maker), and the action can be processed at little or no attentional cost. If, on the other hand, there is a large discrepancy between the priors and the sensory input (e.g., reaching for the teabags and not the coffee), the system has supplied inaccurate predictions, there are many prediction errors, and the action demands allocation of attentional resources. In the no-error case, the action parsing system is stable and can rely on its current action representation to classify the action and predict its time-dependent structure. In the many-errors case, the system is unstable and the current action representation needs to be modified. To modify the action representation, the system has to search through available sensory information for cues that can update the priors. Because the action parsing system has to allocate attentional resources to the available sensory information in order to update, prediction errors are positively associated with external allocation and depletion of attentional resources.

Actions can be parsed on several hierarchically related levels (Hard, Tversky, & Lang, 2006; Zacks, Tversky, & Iyer, 2001). “Lifting a mug from a table, rotating the wrist, pouring a dark liquid substance into the mouth and swallowing” can be finely parsed into four units – “lifting,” “rotating,” “pouring,” and “swallowing” – or coarsely into one unit: “drinking coffee.” While both levels are always available, the default level of action parsing is coarse parsing, which utilizes goal information for the entire action (Sørensen, 2007; Zacks & Tversky, 2001), “drinking coffee” in the example. At this level, prediction errors are goal-driven; they mainly arise if the predicted goal is blocked by the perceived action’s temporal development. Goals are ascribed by priors and the coarse level of action parsing is, therefore, critically dependent on such internal resources. In contrast, the finer level of action parsing – “lifting,” “rotating,” “pouring,” and “swallowing” in the example – is primarily sensory-driven and its predictions are, therefore, influenced by minor variations in the perceived action. In ordinary action perception, these levels are hierarchically aligned in such a way that finer-level units constitute sub-sets of coarse-level units (Kurby & Zacks, 2008).

1.2. Previous studies

Based on the construct of ritualized behavior and insights into how ordinary action perception functions, several empirical studies have, with a few exceptions, found convergent evidence that ritualized behavior has specific perceptual and cognitive effects, and that cultural priors, in the form of socially distributed and experience-dependent priors, can modulate perception and understanding of ritualized behavior.
Keren and colleagues (Keren, Boyer, Mort, & Eilam, 2010) found that ritualized behavior is structured differently to everyday behavioral routines, which are typically processed outside the focus attention. Although both behavioral routines and ritualized behavior are characterized by a high degree of rigidity and repetitions, the latter tends to exaggerate idiosyncratic components and demand increased attention to detail (Boyer & Liénard, 2006). Similarly, two studies have shown that ritualization is a process that diminishes the functionality of instrumental action through repetitions and task-irrelevant action components (Zor, Hermesh et al., 2009; Zor, Keren et al., 2009). In cultural rituals these repetitious action components are often restricted to a very small set of shared actions allowing for considerable behavioral variability (Keren, Fux, Mort, Lawson, & Eilam, 2013).

By modeling ritualized behavior as causally intransparent actions, one study tested the effects of randomizing the sequential order of an action’s motor components on the amount of perceived action boundaries and unit size (Nielbo & Sørensen, 2011). The study also tested the effects of cultural priors, which were operationalized as previous experience with an action, on these variables. Ritualized behavior was consistently parsed into smaller units using more action boundaries than instrumental behavior. Contrary to predictions, there was no indication that cultural priors could modulate parsing of either behavioral type. Because it was possible that the study did not manage to induce priors adequately and that unit size was too coarse a measure to capture modulation effects, Nielbo and colleagues later devised an alternative procedure for induction of priors using conceptual narratives associated with the actions and measured effects of ritualized behavior and cultural priors on hierarchical alignment (Nielbo, Schjoedt, & Sørensen, 2013). Hierarchical alignment is less coarse than unit size and measures the degree to which action boundaries on a fine level of parsing constitute a sub-set of action boundaries on a coarse level of parsing (Zacks et al., 2001). Compared to instrumental behavior, ritualized behavior reliably decreases hierarchical alignment and is, therefore, less hierarchically organized. The study found an effect of cultural priors that, although not statistically significant, indicated how cultural priors can modulate ritualized behavior by increasing perceived hierarchical organization.

A qualitative model of ritualized behavior which combined several of these key findings has previously been proposed (Nielbo & Sørensen, 2013). The model used a continuous signal to track prediction errors during observation of ritualized behavior in an artificial neural network. Simulations showed that while instrumental behavior had only transient increases in the prediction error signal between actions – that is, when one action ends and a new starts – ritualized behavior had a chronically high prediction error signal. This difference occurred, according to the authors, because ritualized behavior forced the system to compute errors at a sensory-driven level instead of a goal-driven level. Since cultural priors can ascribe conventional goals to ritualized behavior, they can make parsing more goal-driven. Further simulations illustrated this claim by showing that cultural priors reliably lowered the prediction error signal during ritualized behavior. Several issues could be raised concerning the model’s validity and scope. First, the model and simulation were overly simplified (Nielbo, Braxton, & Upal, 2012). Actions, for instance, were modeled as dependency relations between simple abstract patterns. Second, the model failed to explain why cultural priors only appear to have a minor, if any, modulation effect in behavior experiments (Nielbo et al., 2013; Nielbo & Sørensen, 2011), while the simulated effect was considerable. Finally, the model did not, for obvious chronological reasons, include recent behavioral insights into ritualized behavior, specifically hierarchical organization (Nielbo et al., 2013).

In the light of previous studies, the present study offers a more detailed and advanced qualitative model of predictions during perception of ritualized behavior. By adding realism and computational complexity, the current model improves Nielbo and Sørensen’s (2013) model significantly in at least four ways: (1) by introducing a new complex learning environment, it shows the robustness of the previous model across instantiations; (2) the new learning environment generates richer and more high-dimensional simulation data; (3) which makes it possible to explain the lack of cultural modulation in experimental studies; and (4) explore new time-dependent non-linear effects of ritualization and cultural modulation.
2. Simulation

The basic system properties needed to model and simulate action parsing and cultural modulation during observation of ritualized behavior are threefold. (1) Action perception: the system must perceive dynamic sequential input from its sensory environment. (2) Perceptual predictions: given sensory input at time $t_1$, the system needs to predict subsequent sensory input at $t_2$. (3) Expectation modulation: the system has to modulate its predictions by experience-dependent information, that is: (3.1) uniquely associated with one specific action only, and (3.2) dissimilar from sensory input. These two information features (3.1 and 3.2) are important to simulate cultural priors that are symbolically mediated. While motor components are shared across different actions, many rituals are, for instance, similar in their combination of approach and reverent motor components; words and concepts can function as unique tags for the type of action they are associated with, for instance “Eucharist” or “Pūjā.” With few exceptions, words are not motivated by the sensory structure of their referent, and a similarity mapping between word and behavior is, therefore, neither necessary nor feasible.

Cultural action rules and norms are not hardcoded in the human mind-brain, but acquired through repeated exposure and instruction (Botvinick & Plaut, 2004; Cooper & Shallice, 2006). To adequately capture this learning aspect of cultural information, it is important that the system learns from experience with actions in its sensory environment. For added realism, it is important that the system’s sensory environment consists of actions that share more than just abstract similarities with human actions. Preferably, the sensory environment should be composed of human actions.

To set up the model, an artificial neural network was used, more specifically an Elman network. An Elman network was chosen for two reasons. First, it has the basic system properties and satisfies the learning constraint. Elman networks learn through experience and are capable of processing sequential information from a dynamic sensory environment. Second, Elman networks have been used to implement previous models of action processing, which makes the present model’s behavior comparable (Botvinick & Plaut, 2004; Hanson & Hanson, 1996; Nielbo & Sørensen, 2013; Reynolds, Zacks, & Braver, 2007). Artificial neural networks (ANNs) are a set of mathematical models that are loosely inspired by the computational structure of biological brains. They consist of computational units (“neurons”) that are linked in layered network structures through weighted connections. This set of models is capable of learning through exposure to an external information environment and subsequent internal processing. An Elman network is a recurrent type of artificial neural network that can learn sequential information, such as actions, by performing one-step predictions. Given an input at one time step, an Elman network can predict input at the following time step. Finally, an Elman network can also use non-distributed “symbolic” information to tag sequential input.

2.1. Technical specifications

2.1.1. Architecture and learning paradigm

Two equations that are typical of abstract ANNs were used to model the sigmoid activation function of the individual units: the net input equation and the activation value equation (O’Reilly & Munakata, 2000). In the model, net input is the sum of weighted input from other units:

$$\eta_j = \sum x_i w_{ij}$$

$\eta_j$ is the net input for the receiving unit $j$, which is a function of the activation value of sending units $x_i$ multiplied by the weight value for that input and the receiving unit $w_{ij}$. The integrate-and-fire model of biological neurons was used to model the activation value of a unit (Abbott, 1999):

$$y_j = \frac{1}{1 + e^{-\eta_j}}$$

$y_j$ is the activation value of the receiving unit $j$.
which outputs the activation value of receiving unit $y_j$ as a function of its net input $\eta_j$.

The basic computational structure of the network consisted of four layers of units, each of which utilized the activation function (see Figure 1): one input layer consisting of 54 units, one hidden layer of 100 units, one context layer of 100 units associated with the hidden layer, and finally one output layer of 54 units. The input layer had full connectivity with the hidden layer; that is, every unit in the input layer was connected to every unit in the hidden layer. Similarly, the hidden layer was fully connected to the output layer. The connectivity between the hidden layer and the context layer used one-to-one projections such that each unit in the hidden layer projected to one, and only one, unit in the context layer. The four layers represented information in a distributed fashion over all the units.

Modulation through cultural priors was implemented by connecting an instructional layer of 24 units to the hidden layer (Botvinick & Plaut, 2004) that was associated with one specific action in the training environment. This layer used a local representation scheme that mimicked symbolic information processing. Each unit in the layer used a linear activation function (specifically the identity function: $y_j = x_j$) and tagged or labeled one action uniquely.

The backpropagation learning algorithm was used to train the networks (Rumelhart, McClelland, & the PDP Research Group, 1986). This is an error optimization algorithm that adjusts the network’s learning (its weight space) over time in order to minimize the difference between the desired learning and actual learning at any time step. Although the algorithm has been criticized for its lack of biological plausibility, it can be understood in a predictive coding framework and interpreted as mirroring the brain’s capacity for error-based learning (Reynolds et al., 2007; Rogers & McClelland, 2008).

2.1.2. Learning environment

To construct a more complex and realistic learning environment than previous studies, the simulation used matrix representations of real instrumental actions performed by a human agent. The matrix representations were based on recordings of a motion-capture system with 18 sensors sampling information from three dimensions giving an 18 by 3 matrix for each motion frame (Reynolds et al., 2007).4 One frame, therefore, consisted of 18 coordinates in a three-dimensional space that characterized the motor component of an action performed by a human agent. An action—that is, a sequence of motor components such as “making coffee”—consisted of a variable number of frames interconnected in a sequential pattern and presented step-wise to the networks. One frame
represents one specific motor component ("lifting the kettle" or "turning on the water"). Given a motor component input ("turning on the water"), the networks’ task was to predict the subsequent motor component ("filling the kettle").

Ritualized actions were modeled using a two-step procedure, by randomizing the order of motor components and using an unequal distribution of actions respectively. Both randomization of order and unequal distribution of actions represent central assumptions about the nature of actions and action perception in the model. At a formal level, it follows from these two assumptions that in the model ritualized actions are essentially low-frequency actions that share motor components with high-frequency actions.

First, to make the simulation similar to previous studies (Nielbo et al., 2013; Nielbo & Sørensen, 2011, 2013), ritualized actions were modeled by randomly reorganizing the motor components of the instrumental actions.5 If, for instance, the instrumental action of “making coffee” consisted of the sequence of motor components “lifting the kettle” → “turning on the water” → “filling the kettle,” then a ritualized counterpart might be “lifting the kettle” → “filling the kettle” → “turning on the water.” With this transformation we wanted to model ritualized actions as causally intransparent actions that share motor components with causally transparent actions. The causal intransparency of a ritualized action is, to a large extent, a product of the recombination of the motor components that are normally used in the context of instrumental actions, which divorces the motor components from their typical instrumental goals. The Eucharist, for instance, use several motor components that are typically related to feeding behavior (eating bread and drinking wine or grape juice), but the motor components are no longer related to their instrumental goals (replenishment of calories and rehydration). Although the ritualized action ascribes new symbolic meanings to the motor components, these meanings are still constrained by the instrumental goals of the motor patterns (participants in the Holy Communion eat the body and drink the blood of Christ).

Second, to implement a preference for causally transparent actions, we set up a training environment with an unequal distribution of actions types, that is, instrumental actions were more frequent than ritualized actions. The networks would therefore tend to expect that a sequence of motor components look like instrumental actions (i.e., causally transparent), but they would also have learned that under certain conditions the same motor components would be organized differently and look like ritualized actions (i.e., causally intransparent). It is important to notice that before training, the networks have no particular “innate” sequential preferences and that any sequence of motor patterns therefore is equally (in-)transparent. The current training environment consisted of approximately 30% ritualized actions and 70% instrumental actions. Although the specific distribution of ritualized actions in a person’s environment will vary depending on context and profession (e.g., both the Adhvaryu priest and OCD patient experience more ritualized behaviors than the average person), we believe that this estimate is actually quite high compared to what we expect a person to experience during an extended time period. For logical reasons we assume that the instrumentality of an action (its degree of causal transparency) is positively associated with its frequency, but do acknowledge that the actual distribution is ultimately an empirical question.

To associate cultural priors with specific actions, a unit in the instructional layer would fire maximally during the full sequential presentation of its associated action and otherwise not fire at all. The networks could, therefore, use the activation of this layer to determine which action was presented.

The training and testing set then consisted of two sub-sets, an instrumental and a ritualized, within which each action in one sub-set had a counterpart in the other sub-set. For instance, the instrumental sub-set contained a “having a drink” action pattern which had a ritualized counterpart in the ritualized sub-set, namely a randomly reorganized version of “having a drink.” The total number of action patterns in the training and testing set was 24 (see Table 1).

2.1.3. Measures
Since the goal is to model action processing as a prediction process using a continuous prediction error signal – that is, a fine-grained measure of the networks’ accuracy in predicting the correct
motor components in a sequence – the output layer’s summed squared error (SSE), which measures the summed difference between a predictive target and an actual output, was chosen as an appropriate measure of prediction error on training and test trials:

\[ \text{SSE} = \sum_k (t_k - o_k)^2 \]

where \( t_k \) is the target output and \( o_k \) the actual output of the \( k \)th unit.

2.1.4. Design and procedure

The simulation was set up as a mixed 2×2 factorial design with action type (instrumental/ritualized) as the within-subjects factor and expectation modulation (control/cultural priors) as the between-subjects factor. SSE was consistently used as the dependent variable.

To investigate if the networks exhibited hierarchical organization in their action processing, we used a 2×2 within-subjects factorial design, with transitions (between/within) as the first factor and action type (instrumental/ritualized) as the second factor. The logic behind the first factor was that if the networks showed a clear differentiation in SSE when it came to transitions between actions (i.e., boundaries between the final sub-action in one action sequence and the initial sub-action in another action sequence) and transitions within an action (i.e., boundaries between sub-actions in one action sequence), then they distinguished between different hierarchically related levels of action processing.

To implement both designs, 40 networks (\( N = 40 \)) were trained and tested on the entire training set of 2×12 action patterns. Of these networks, 20 (\( n_t = 20 \)) were trained with cultural priors (i.e., with activation of the instructional layer), while 20 (\( n_c = 20 \) networks were trained as controls (i.e., with a disconnected instructional layer). The networks were trained to asymptote over 2000 epochs (one epoch equals one presentation of the full stimulus set) and subsequently tested once on the entire set of action patterns. The Emergent Neural Network Simulation System was used to implement the simulation the model’s behavior (Aisa, Mingus, & O’Reilly, 2008).

2.1.5. Analysis

We used an \( \alpha \)-level of .05 for tests of statistical significance, and \( p \)-values are only reported if \( F > 1 \). When needed, the Bonferroni-Holm sequentially rejective multiple test with a family-wise \( \alpha \)-level of .01 was used to correct \( p \)-values for multiple comparisons (Holm, 1979). Eta-squared (\( \eta^2 \)) was used for effect size measures. The Shapiro-Wilk \( W \) test (Shapiro & Wilk, 1965) was used to confirm that the data did not deviate significantly from normality.

<table>
<thead>
<tr>
<th>Number</th>
<th>Action sequence</th>
<th>Number of frames</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Bowing</td>
<td>10</td>
</tr>
<tr>
<td>2.</td>
<td>Cheering</td>
<td>10</td>
</tr>
<tr>
<td>3.</td>
<td>Having a drink</td>
<td>12</td>
</tr>
<tr>
<td>4.</td>
<td>Driving</td>
<td>11</td>
</tr>
<tr>
<td>5.</td>
<td>Opening a door</td>
<td>9</td>
</tr>
<tr>
<td>6.</td>
<td>Sawing</td>
<td>12</td>
</tr>
<tr>
<td>7.</td>
<td>Sitting down</td>
<td>11</td>
</tr>
<tr>
<td>8.</td>
<td>Using a sledge hammer</td>
<td>10</td>
</tr>
<tr>
<td>9.</td>
<td>Spilling a drink</td>
<td>11</td>
</tr>
<tr>
<td>10.</td>
<td>Standing up</td>
<td>11</td>
</tr>
<tr>
<td>11.</td>
<td>Swatting bees</td>
<td>12</td>
</tr>
<tr>
<td>12.</td>
<td>Having a tantrum</td>
<td>10</td>
</tr>
</tbody>
</table>

Table 1. The 12 actions that were used in the instrumental and the ritualized sub-set during training and testing of the model.
3. Results

3.1. Error signal, cultural priors, and hierarchical organization

Mean SSEs for all the networks were examined to determine how ritualized behavior, cultural priors, and hierarchical organization influenced their error signals during the test trials (Figures 2 and 3).

A 2 (action type: instrumental/ritualized behavior) × 2 (priors: cultural prior/no cultural prior) × 2 (event boundary: between actions/within action) mixed model ANOVA was conducted to analyze the error signals. There was a main effect of action type, \( F(1, 19) = 232.08, p < .0001, \eta^2 = .93 \), which indicated a large difference in error signal between action type conditions. Analysis also showed a medium main but significant effect of prior, \( F(1, 38) = 6.84, p < .05, \eta^2 = .31 \), and finally a very strong main effect of event boundary: \( F(1, 19) = 4286.65, p < .0001, \eta^2 = .98 \). There was no significant action type by prior interaction, \( F(1, 38) = 3.3, p = .08, \eta^2 = .17 \), but a significant action type by event boundary interaction, \( F(1, 19) = 218.69, p < .0001, \eta^2 = .76 \), as well as a prior by event boundary interaction, \( F(1, 38) = 35.13, p < .0001, \eta^2 = .31 \). Post hoc tests corrected for multiple comparisons supported these observations by showing reliable difference between action type, priors, and event boundary. The two-way interactions can be explained by a statistically reliable three-way interaction, \( F(1, 38) = 30.06, p < .0001, \eta^2 = .31 \), which shows that priors have a differential effect on the action by event boundary interaction, more specifically, priors lower the within action boundary prediction error signals for ritualized action more strongly than for instrumental behavior. It can be observed in Figure 2(right panel) that for ritualized behavior, cultural priors decrease the error signal within an action, while the error signal between actions is virtually unchanged. Furthermore, the relative difference in error signals for event boundaries between actions and event boundaries within an action is larger for ritualized behavior than for instrumental behavior.

Figure 2. Mean error signal (SSE) for both action types with and without cultural priors (left) and mean error signals on event boundaries between actions and within actions (right). Error bars indicate 95% confidence interval.
3.2. Synchronous error signal

To explore why recent behavioral studies have not managed to track an effect of cultural priors, the four time series presented in Figure 3 were correlated. The reasoning behind these analyses was that if the predicted modulation effect exists, it is likely that the measurement instrument in previous experiments lacks sensitivity. Since the behavioral experiments used discrete button presses as a proxy for prediction errors, we reasoned that cultural priors do not modify the frequencies of the error signals (i.e., the number of event boundaries), but rather the amplitudes of the error signals (the magnitude of the event boundaries). This account of the effect of cultural priors would mean that error signals should be synchronous and therefore correlated between the two prior conditions (cultural prior/no cultural prior). This was confirmed by respective correlation coefficients: mean SSEs for instrumental behavior with and without cultural priors were significantly correlated, $r = .9, p < .01$, as were mean SSEs for ritualized behavior, $r = .69, p < .01$. This is not particularly surprising, since the error signals follow the general hierarchical organization of the actions, indicated by the iterated valley-peak structure in each time series of Figure 3. The behavioral experiments, however, only measured event boundaries within actions; we therefore calculated the correlation coefficients for the event boundaries within the actions. The analyses revealed a similar but weaker correlation pattern: SSEs for both instrumental, $r = .35, p < .01$, and ritualized, $r = .2, p < .05$, behavior with and without cultural priors were reliably correlated.

**Figure 3.** Error signal time series for both action types with and without cultural priors. The recurrent valley-peak structure indicates the hierarchical nature of each signal, with valleys indicating event boundaries within an action and peaks indicating event boundaries between actions (punctured vertical lines).
3.3. Dynamics of cultural priors

To further explore the model’s dynamics, we performed Recurrence Quantification Analysis (RQA) on the four time series for each network. RQA is a non-linear method for analyzing dynamical systems and their trajectories, which can capture dynamic properties that are lost when averaging in standard correlational methods (Marwan, Romano, Thiel, & Kurths, 2007; Riley & Van Orden, 2005; Webber, 2012).

Initially, we searched for qualitative patterns of recurrence in the error signals using plots of the individual time series (Figure 4). A recurrence plot is a two-dimensional plot that visualizes recurrences in a time series by plotting the time series along both the abscissa and the ordinal. Visual inspection of the plots indicated different global patterns of organization in the contrast between networks with and networks without cultural priors. Networks with cultural priors had a dense global pattern of recurrence. Long periods of recurrence (long diagonal lines) and long vertical lines indicate that the system is paused in a singularity. Both properties could be mapped onto some or all of the 12 individual actions in the time series. These recurrence patterns suggest long periods of stability within each action and a high degree of periodicity in the systems with cultural priors. In contrast,
the recurrence patterns of networks without cultural priors were sparser and only showed short periods of recurrence (short diagonal lines), indicating more chaotic dynamics. To further qualify these observations, we used two RQA metrics that quantify the recurrences of the plots (Figure 5). These metrics were %determinism (%DET), which measures the predictability of the plot as the percentage of recurrent point that forms diagonal lines and %laminarity (%LAM) that measures the smoothness (as opposed to chaoticity) of the plot as the proportion of recurrent point that forms vertical lines.

Based on the error signal recurrences plots, we computed the two metrics and compared them in a 2×2 mixed-model multivariate analysis of variance (MANOVA), with prior (cultural prior/no cultural prior) as the between-subjects variable and action type (instrumental/ritualized behavior) as the within-subjects variable (see Figure 5 for descriptive statistics and Table 2 for significance values). Using Pillai’s trace, there was a significant main effect of prior, $V = 0.22, F(2,75) = 10.95, p < .0001$, across both metrics; a significant main effect of action type, $V = 0.15, F(2,75) = 6.51, p < .01$, approaching significance for LMAX; and no significant prior by action type interaction, $V = 0.05, F(2,75) = 2, p = .14$. From these comparisons we can infer that cultural priors reliably make action processing more predictable, stable, and smooth, which supports the observation of periodic, as opposed to chaotic, dynamics for networks with cultural priors.

### 3.4. Discussion

Analyses confirmed that the individual networks did, indeed, show increases in their error signal when processing ritualized behavior compared to instrumental behavior. The model hereby behaves in accordance with previous behavioral experiments in which participants placed more event boundaries during segmentation of ritualized actions in contrast to instrumental actions. In contrast to the

![Figure 5](image-url). Mean of the two RQA metrics; left %determinism (%DET) and right %laminarity (%LAM) for both action types (instrumental/ritual) with and without cultural priors (no prior/prior). Error bars indicate 95% confidence interval.
behavioral experiments, the simulation analyses showed that cultural priors effectively modulated the error signal, especially by decreasing the signal’s magnitude during ritualized behavior.

Similarly, the model was capable of simulating the presence of, as well as a difference in, hierarchical structure during processing of both ritualized and instrumental behavior. Interestingly, cultural priors affected ritualized behavior more strongly than instrumental behavior by increasing the hierarchical structure. Ritualized behavior, therefore, seems particularly sensitive to the modulation effect of cultural priors. While the previous model (Nielbo & Sørensen, 2013) has not been able to explain the absent effect of cultural priors in behavioral experiments, the present model showed that this absence originates in synchronous error signals. We all utilize the same cues to segment instrumental and ritualized behavior, and our error signals are therefore correlated independently of cultural priors. The magnitude of the error signal, however, differs as a function of cultural priors. A cultural expert (with cultural priors) and the cultural novice (without cultural priors) both utilize the available psycho-physic information to segment an action, but their priors modulate the error signal’s order of magnitude differentially. When, for instance, a Christian priest observes a baptism, s/he uses the same physical discontinuities and relations to segment the action as does the initiate’s godmother. Their error signals, therefore, show a similar frequency. Having cultural priors readily available from multiple observations, discussions, and teachings of the baptism, the priest, compared to the godmother, is less “surprised” by the unfolding of the action. The magnitude of the priest’s error signal is, therefore, of a smaller order than the godmother’s.

Finally, through non-linear tools of analysis, the model indicated that cultural priors increase system stability and periodicity. Together with the general reduction of the error signal by cultural priors, these findings support the idea that cultural information can increase stability of psychological systems and that there might be a pressure on constructing such information when socially sanctioned behavior, such as ritual, tends to decrease system stability (Nielbo & Sørensen, 2013).

### 4. General discussion

In this article, we have advanced a qualitative model that accounts for the action perception and cultural modulation of ritualized behavior by combining three computational resources: sequential sensory input, priors, and error monitration. We simulated the model’s behavior using a more complex and realistic learning environment and network architecture than previous models. Through simulations we have shown that the model can account for (1) transient increases in error signals during perception of ritualized behavior and modulation effects; and (2) modulation of these error signals by cultural priors.

Computational models are, as with any other models, simplified versions of the systems they represent (Nielbo et al., 2012). Two such simplifications stand out in the present model. First, it is a model of ritualized behavior, a scientific construct that is characterized by a specific feature set and which is often embedded in or part of the many diverse behaviors that are labeled rituals. Therefore, we do not claim that the model explains ritual as such, but only specific perceptual and cognitive effects that are related to causally intransparent and goal-demoted actions. Second, in contrast to a

<table>
<thead>
<tr>
<th>Effect</th>
<th>Metric</th>
<th>F</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Between subjects</td>
<td>%DET</td>
<td>15.88</td>
<td>0.000*</td>
</tr>
<tr>
<td></td>
<td>%LAM</td>
<td>19.90</td>
<td>0.000*</td>
</tr>
<tr>
<td>Within subjects</td>
<td>%DET</td>
<td>23</td>
<td>0.633</td>
</tr>
<tr>
<td></td>
<td>%LAM</td>
<td>2.00</td>
<td>0.165</td>
</tr>
<tr>
<td>Prior × action type</td>
<td>%DET</td>
<td>1.77</td>
<td>0.188</td>
</tr>
<tr>
<td></td>
<td>%LAM</td>
<td>2.93</td>
<td>0.091</td>
</tr>
</tbody>
</table>

*Statistically significant.
recent agent-based hybrid model of ritual transmission (Whitehouse, Kahn, Hochberg, & Bryson, 2012), we have omitted affective and motivational components. This is not because these components are irrelevant to ritualized behavior, as illustrated by Boyer and Liénard’s Hazard-Precaution model (2006), but given previous studies (Keren et al., 2010; Nielbo et al., 2013; Nielbo & Sørensen, 2011, 2013; Zor, Hermesh et al., 2009) and the model’s perceptual and cognitive scope, it was a necessary omission for model interpretability.

4.1. Error, attention, and priors

In line with previous results, the simulation unanimously showed that ritualized behavior increased prediction error signals. This indicates that the increase in error signals makes a perceptual system unstable because it allocates all its attentional resources to the immediate sensory environment, computing minor variation in behavioral trajectories and thereby depleting perceptual and cognitive resources. These attentional resource-related effects explain why ritualized behavior is experienced as attention demanding and salient with a focus on minute and goal-irrelevant detail. At the level of categorization, differences in a continuous error signal are likely to elicit different modes of categorization, such that actions associated with high levels of error (“ritual”) are categorized as distinct from actions associated with low levels of error (“instrumental”). The simulations, therefore, support the claim that ritualized behavior embedded in rituals is a generative factor for the, more or less, universal distinction between instrumental and ritual behaviors.

How then does a system handle the instability related to ritualized behavior, and how does it distinguish between ritualized behavior embedded in socio-cultural ceremonies and their pathological counterparts? The system utilizes internal resources, typically socially shared information associated with the ritualized behavior (i.e., cultural priors), which modulate the error signal, change the chaotic dynamics to periodic, and thereby stabilize the system. Ritualized behaviors embedded in religious rituals are often represented as culturally meaningful because they evoke shared expectations in participants and because they are associated with specific, although abstract, goals. In the pathological case, the system stability is not supported by such priors, but it is likely that the system tries to construct individual priors ad hoc to emulate the effects of cultural priors. The model thus points to a possible link between cultural narratives that are concerned with rituals and the ritual performances. Since cultural priors can stabilize perceptual representations of ritualized behavior, ritual performances might actually be a key motivating factor in the construction of cultural narratives. One can only speculate with regard to the model’s consequences for the relation between compulsive behavior and obsessive thoughts.

From an error optimization perspective, an ironic consequence of how cultural priors modulate action processing is that priors can convert ritualized behavior, which demands a lot of attention, into routinized behavior, which demands little or no attention (Keren et al., 2010). More evolutionary-oriented models are likely to find this problematic and therefore include affective and motivational components to retain a fundamental difference (Boyer & Liénard, 2006). While differences in motivation could change this outcome, it is important to notice that although cultural priors did decrease the error signal and stabilize the system into more periodic dynamics, we still found a reliable difference in error signals between the action type conditions. In other words, cultural priors reduce error signals elicited by ritualized behavior; they do not erase them.

4.2. Cognitive approaches to ritual

In modeling how ritualized behavior is processed in terms of action perception and cognition, several aspects of the model are relevant to some of the recent developments in cognitive approaches to ritual.

According to Divergent Modes of Religiosity (DMR), social ritual systems tend toward either a doctrinal or an imagistic attractor (Whitehouse, 2002; Whitehouse et al., 2012). At the level of
cognition, the attractors differ with respect to the type of memory utilized for encoding and transmission. We believe that the interplay between features of a behavior and cultural priors might influence mnemonic strategy. As discussed elsewhere, a tradition that embeds highly resource-depleting behaviors in rituals will effectively block online memory encoding (Schjoedt et al., 2013). The attractor differences in our model are best understood as a difference in terms of the content of cultural priors. The difference between imagistic and doctrinal priors should then be a gradual transition from high to low levels of experiential and individualized content. Memory-related effects originating in content-specific modulation of error signals could be one interesting developmental trajectory for the model.

At the level of ritualized behavior, one interesting challenge to the model is DMR’s proposed tedium effect, which states that frequency and motivation of ritual performance are inversely correlated (Whitehouse et al., 2012). When ritual performance is very frequent, the social ritual system runs the risk of saturating participant motivation. Conceptually, our model’s counterpart is sensory adaptation through extensive exposure to a specific set of ritualized behaviors. To effectively reduce the error signal by sensory adaptation, the system would have to learn from an environment where ritualized behaviors were considerably more frequent than instrumental behaviors. The system’s error signals would then slowly adapt to the behavioral distributions. We have not implemented the scenario, which in real-world terms would be comparable to performing the Eucharist multiple times a day.

Findings from development psychology indicate that causal opacity in ritualized behavior can function as a social learning signal, which increases imitative fidelity in children (Herrmann, Legare, Harris, & Whitehouse, 2013). Causal opacity activates a ritual stance, in contrast to an instrumental stance, which orients the learner toward an action rationale based on cultural conventions rather than physical causation (Legare & Herrmann, 2013). According to our model, low-level action integration and non-periodicity of the error signal are automatic and cannot be avoided in ritualized behavior. It is highly likely that social learning systems utilize stable increases in error signals to activate certain hyper-priors comparable to a ritual or an instrumental stance, which in the ritual case reduce the error signal.

Another related line of research has argued that ritual can serve as a mechanism for coping with affective responses elicited by randomness and lack of control. One study in particular found that priming randomness increased perceived efficacy of ritual (Legare & Souza, 2014). One possible way the model of ritualized behavior can contribute to this argument is through temporary thought suppression. By depleting attentional resources, ritualized behavior can effectively block appraisals of affective states independently of the eliciting factors. As Boyer and Liénard have pointed out, the irony of this mechanism is that over time it will tend to increase the need for thought suppression and thereby reinforce ritualized behavior (Boyer & Liénard, 2006).

One final issue needs to be addressed, namely the distinction between models of proximate causation (i.e., how a mechanism works) and models of ultimate causation (i.e., why a mechanism exists). The present model only solves problems of proximate causation, specifically how ritualized behavior is processed at the level of perception and cognition, and we have therefore not approached questions such as why ritualized behaviors are so persistent in our species’ behavioral matrix. Several ultimate causation models of ritual and ritualized behavior already exist (e.g., Boyer & Liénard, 2006; Bulbulia & Sosis, 2011), which our error-based accounts are compatible with (Schjoedt et al., 2013).

According to cooperative signaling theories of ritual, social rituals facilitate communication and cooperation by signaling cooperation-relevant properties through honest or costly displays (Alcorta & Sosis, 2005; Bulbulia, 2004; Bulbulia & Sosis, 2011; Sosis, 2004). While recognition of honest displays is likely to have innate constraints, their saliency can, at least in part, be a direct function of embedded ritualized behavior and related error signals. Our model, therefore, offers a proximate mechanism that can explain the perceptual and cognitive underpinnings of cooperative signaling theories in humans.
Boyer and Liénard have offered an ultimate model in which ritualized behavior is a response to the activation of an evolved hazard-precaution system. According to their model, ritualized behavior swamps working memory temporarily and suppresses intruding thoughts during system calibration and misfiring. Part of the model contains an action parsing system, which responds to ritualized behavior by directing attention to low-level detail of the stimulus. We have computationally implemented the action parsing system based on specific architectural assumption (i.e., sensory input, priors, and error monitoration). In our model, swamping is caused by a large error signal that results in resource depletion. Suppression of intruding thoughts can therefore be understood as a specific consequence of a general mechanism.

From a proximate perspective, our model offers a link between individual and pathological ritualized behaviors on the one hand and socially sanctioned communal group displays through the error-reducing properties of cultural priors on the other hand. Features of ritualized behavior are always attention demanding, but they are less depleting for groups that share cultural narratives than for individuals who use the behavior to suppress intruding thoughts.

Notes

1. By cultural priors we mean experience-dependent information that is socially distributed and which functions as a stable prediction resource in action perception.
2. The action parsing system tracks information on multiple levels. “Drinking coffee” can, for instance, be embedded in “becoming a barista” which then again can be embedded in “attaining a happy life” (Zacks & Tversky, 2001). For the sake of simplicity, only the basic level and the lower level are implemented in the current model.
3. A thorough description of artificial neural networks and Elman networks is beyond the scope of this article. For a conceptual or formal introduction, respectively, see Gurney (1997) and Hagan, Demuth, and Beale (2002).
4. We are most thankful to J.R. Reynolds, J.M. Zacks and T.S. Braver for making a coded motor library available and furthermore to J.M. Zacks for suggesting possible implementations. See http://dcl.wustl.edu/DCL/Stimuli.html (3 October 2011) (Reynolds et al., 2007).
5. We did not implement redundancy by repeating motor components in the ritualized actions for three reasons: (1) redundancy would increase the error signal in the ritualized condition because the network is sensitive to the total number of motor components; (2) the differing number of motor components between conditions would make the conditions less comparable; and (3) this would also make results less comparable to previous studies where insertion of additional motor components has been avoided for methodological reasons (Nielbo et al., 2013).

Acknowledgments

The authors would like to thank two anonymous reviewers for their comments on a previous version of the article.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Danish Ministry of Science, Technology and Innovation [UNIK grant MINDLab] and the Canadian Social Sciences and Humanities Research Council [grant Cultural Evolution of Religion Research Consortium].

ORCID

Kristoffer L. Nielbo © http://orcid.org/0000-0002-5116-5070
References


