

# **MATERIAL CORRELATES ANALYSIS (MCA): AN INNOVATIVE WAY OF EXAMINING QUESTIONS IN ARCHAEOLOGY USING ETHNOGRAPHIC DATA**

In the past decade, a number of large, longitudinal cross-contextual historical and ethnographic databases have been used to test theories of group dynamics (Atkinson and Whitehouse 2011; Johnson 2005; Peoples and Marlowe 2012; Watts et al. 2015). In spite of their potential impact on our understanding of prehistory, some of these theories have never been tested using archaeological data or only partially (e.g. Mithen 2004; Whitehouse and Hodder 2010). However, the emergence of large cross-cultural archaeological and ethnographic databases (e.g. Electronic Human Relations Area Files), as well as the application of multivariate statistical and network analysis techniques now offer the opportunity to integrate data from ethnographic, historical, and archaeological resources into a single diagnostic framework for the purpose of hypothesis testing. These approaches bring us closer to examining and validating claims regarding the dynamics of human groups in prehistoric contexts. In this paper we outline a new method of Material Correlates Analysis (MCA) which includes targeted data gathering and use of a specific set of statistical techniques for the purpose of hypothesis testing using ethnographic and archaeological material.

In our example, we use a theory developed in the cognitive anthropology of religion—the Modes of Religiosity theory (Whitehouse 1995)—as a framework to assess the causal link between social bonding through rituals and the generation of the required levels of group cooperation in emerging sedentary agriculturalist communities, during the initial agricultural transition in south-west Asia. The modes theory describes different forms of group or social bonding based on a distinction between low frequency, high arousal experiences—classed as imagistic rituals and high frequency, low arousal experiences—classed as doctrinal rituals (Whitehouse 1995). Pertinent to our research, recent analysis of large datasets of rituals from across the globe demonstrates that agricultural societies appear to be characterized by a tendency toward the doctrinal mode of religiosity (Atkinson and Whitehouse 2011; Whitehouse and Hodder 2010; Whitehouse et al. 2014). In spite of a number of promising qualitative

observations, the full potential of the modes theory to examine social cohesion in the archaeological record has not been utilized in any systematic manner— particularly, in relation to the agricultural transition—as any examination of the modes using archaeological data inherently has one major problem: that ritual frequency and arousal levels are often impossible to discern archaeologically with confidence due to material culture gaps in the archaeological record. Typically, ritual activity and religious beliefs can only be inferred indirectly from the archaeological record, making results less certain than the documented evidence available for the ethnographic samples. To limit the effect of this problem, the MCA method we outline uses sets of relevant material correlates derived from ethnographically known imagistic and doctrinal cultures, previously classified as imagistic or doctrinal by religiosity Atkinson and Whitehouse (2011), to examine (inferentially) the presence of imagistic or doctrinal modes of social cohesion in otherwise scant evidence from prehistory.

In what follows we use the Modes of Religiosity theory to illustrate the application of MCA for the purpose of hypothesis testing using a combination of ethnographic and archaeological data. By creating a generalized linear modelling (GLM) matrix combining the ethnographic and the archaeological data, we demonstrate how sets of ritual and non-ritual material correlates can be used to generate a percentage probability that shows how archaeological evidence indicative of either the imagistic or doctrinal mode can be used to classify each archaeological case. In doing so, we examine one hypothesis—namely, that the agricultural transition is connected with the emergence of a more doctrinal mode to achieve the required levels of group cohesion in emerging sedentary agriculturalist communities. In presenting our example we (1) outline the application of the MCA method to examine a theory developed and tested by means of anthropological data using archaeological material culture and (2) make the initial steps in expanding the reach of the modes theory beyond the currently available historical and ethnographic sources, to test certain hypotheses longitudinally in relation to the prehistory.

### *Modes of Religiosity*

The Modes of Religiosity theory (Whitehouse 1995) proposes two principal courses for the transmission of religious ideas via ritual activity: the imagistic mode and the doctrinal mode. Imagistic ritual practice (low frequency, high arousal) involves the transmission of religious ideas through irregularly occurring, highly emotive experiences stored and recalled via episodic memories that specify who else was present during a given collective ritual performance (Whitehouse 1995, 2000). They are associated with highly cohesive local groups such as bands or tribes that rely on high levels of dependence between group members. In contrast, the doctrinal mode (high frequency, low arousal) centres on the transmission of religious ideas through frequent repetition of highly prescribed ritual practices which are stored in an individuals' semantic memory specifying generic roles rather than individuated participants. The doctrinal mode is associated with larger (often) more geographically extensive groups. The doctrinal mode is considered a more recent development connected with the emergence of more centralized societies (Whitehouse 2004; Whitehouse and Hodder 2010).

To date, the modes theory has been readily used as a methodological framework in anthropological, historical, and cognitive science research using a variety of historical and ethnographic sources (Atkinson and Whitehouse 2011; Malley 2004; Mithen 2004; Naumescu 2008; Pachis and Martin 2009; Whitehouse 2002, 2004; Whitehouse and Hodder 2010; Whitehouse et al. 2014; Xygalatas 2012). Directly relevant to our research focusing on the agricultural transition is the extensive ethnographic examination of group cohesion via ritual activity by Atkinson and Whitehouse (2011). Analyzing data on frequency, arousal level and social structure for 645 religious rituals from a global, cross-cultural sample of 74 cultures via the electronic Human Relations Area Files (eHRAF), they statistically demonstrate that agricultural 'intensity' is a significant predictor of mode of religiosity—that ritual frequency correlates positively with agricultural intensity and that dysphoric intensity correlates negatively with agricultural intensity. They define agricultural intensity on a scale from no reliance on agriculture to full dependence on agriculture (including cultivated crops and herded animals). This study suggests that agricultural activity may have been responsible for a general

increase in the frequency of communal rituals and indirectly presented opportunities for other features of the doctrinal mode to appear—for example, the formation of uniform and prescribed regional traditions. Extrapolating from the ethnographic research by Atkinson and Whitehouse (2011), we use the MCA method to examine the connection (if any) between group cooperation during the initial agricultural transition in south-west Asia and the emergence of the doctrinal mode of religiosity.

### ***Analogy, Ethnoarchaeology, and the MCA approach***

Since the 19<sup>th</sup> century, archaeologists have used some form of ethnographic analogy to interpret archaeological material (David and Kramer 2001; Sillar and Joffré 2016; Stiles 1977). In the context of the *new* or processual archaeology of the 1960s, ethnoarchaeology gained prominence as sub-discipline of archaeology. In one of the formative papers in the development of ethnoarchaeology, Ascher (1961) discusses a number of the essential theoretical and methodological issues associated with the use of ethnographic analogy to infer past behaviour. Middle Range Theory (Binford 1967, 1978)—which advocates the use of analogical inference and the objective testing of hypotheses to examine connections between the present and the past—greatly influenced the development of ethnoarchaeology. A principal objective of ethnoarchaeology has been to move away from *simple* suppositions of cultural continuity and to establish a more comprehensive approach to interpretation by identifying predictable features of human behavior (David and Kramer 2001; Sillar and Joffré 2016; Stiles 1977). Since the 1960s, a number of cross-cultural ethnographic analogies have been utilized to interpret individual objects, techniques (production techniques and use), as well as wider issues of social and economic organization such as exchange networks and social hierarchies (Sillar and Joffré 2016; Stiles 1977).

Since the 1960s, the validity of ethnographic analogy has been heavily debated (Ascher 1961; Binford 1967, 1978; Fewster 2006; Gould 1980; Gould and Watson 1982; Hodder 1982, 1986; Lane 1995; Orme 1973, 1974; Oswalt 1974; Politis 2015; Ravn 2011; Shelley 1999; Stiles 1977; Trigger 1989; Watson 1979; Wobst 1978; Wylie 1985, Yellon 1974). For example, Wobst (1978) states that relying on ethnographic analogy limits us to interpreting the past based on behaviours accessible only via

(current) ethnographic data—described as the “tyranny” of ethnography. Hodder (1982) points out that when employing ethnographic analogy, we must be aware of the inherent subjectivity in using present ethnographic data to interpret the past—making it difficult to use analogy to make valid inference regarding archaeological material.

A number of researchers, critical of the problematic assumptions associated with ethnographic analogy, advocate a more systematic use of cross-cultural ethnographic data in archaeological analysis. For example, building on McNett (1979) and Murdock’s (1957) ethnological approach, Ember and Ember (1995) advocate the use of ethnographically discerned material correlates or proxy measures of human behaviours which can be used to examine statistically (via covariance) both causal (direct) and non-causal (indirect) links between variables (Ember and Ember 1995; Peregrine 1996). In doing so, they demonstrate how a systematic material correlates approach has the potential to aid our interpretation of the archaeological record.

Similarly, Ensor (2003; 2011; 2017) outlines how taking a cross-cultural ethnological approach using material correlates—focusing on evidence for changes in resources and production—presents archaeologists with a framework to examine social transformation in prehistory via empirical archaeological inference. Peregrine (1996; 2001) asserts that results generated from detailed cross-cultural research may represent an appropriate source for generating statistically valid inferences to identify and examine behavioural trends and associations. For example, Peregrine (1993, 1994, 1996) suggests that settlement patterns and house forms reflect identifiable aspects of material culture that can be readily used in a systematic material correlates approach

Of direct relevance to our study, Peregrine promotes the use of cross-cultural ethnographic databases such as the Human Relations Area Files (HRAF) (1) to study the causal and non-causal associations between sets of material correlates and (2) to develop sets of correlates to examine non-material aspects of prehistoric culture such as religious beliefs—both greatly enhancing our interpretation and understanding prehistoric cultures (Peregrine 1996, 2001).

The traditional use of analogy has its limitations, particularly at the larger-scale cultural level. Building on the work of Ember and Ember, Ensor, and Peregrine, we developed a systematic cross-cultural approach (MCA), which uses particular sets of material correlates and statistical modelling to test our hypothesis.

## **METHODOLOGY**

### ***Sample Selection***

To test our hypothesis it was necessary to design an integrated data gathering and analysis framework to collect, categorize and quantify data from the eHRAF ethnographic database and the available archaeological sources. The known eHRAF cultures in our research were a defined subset of the original 74 groups classified as imagistic or doctrinal by Atkinson and Whitehouse (2011). Two criteria were used to select the subset of cultures: the dysphoric arousal levels (average dysphoric mean) and ritual frequency (frequency per year). The rationale was to generate an ethnographic sample that would relate directly to the central aspects of the Modes of Religiosity theory—ritual frequency and arousal level. Of the 74 cultures, those that represented the “most” imagistic and doctrinal cultures were selected, producing a subset of 34 cultures: 15 imagistic and 19 doctrinal. In order to examine the archaeology systematically, data relating to 50 site-phases from across the agricultural transition in south-west Asia were assembled representing the Epipalaeolithic to the Pottery Neolithic cultural horizons. To control for potential biases and limitations associated with cross-cultural studies, the recommendations by Ember and Ember (2009) and Levinson and Malone (2000) regarding targeted data recording and analysis were employed when generating the data for this research.

We identified and recorded sets of material correlate data from the known sample of 34 eHRAF doctrinal and imagistic cultures which were used (1) to identify and record sets ritual (apart from ritual frequency and arousal level) and non-ritual material correlate variables that were directly connected to the known (imagistic or doctrinal) cultures in the ethnographic sample and (2) to derive material correlate variables that could be used to explore the presence of the imagistic or doctrinal

mode in the archaeological samples. In total, we identified a set of 90 ritual, subsistence, and social complexity material correlate variables that could be examined in both the ethnographic and archaeological record (Supplemental Text 1).

### ***Ethnographic Archaeological Data Gathering and Categorizing***

The ethnographic component of this research focused on the collection, cataloguing and analysis of a set of ritual, subsistence practice, and social complexity material correlate variables from the sample of 34 previously classified ethnographic cultures provided by the eHRAF (electronic Human Relations Files) cross-cultural database (<http://ehrafworldcultures.yale.edu/ehrafe/>). Outline of Cultural Material (OCM) codes—a specific set of search codes used to catalogue and search the cultural information provided by the eHRAF database—were used to extract and record the information relating to each of the selected ethnographic cultures (<http://hraf.yale.edu/resources/reference/outline-of-cultural-materials/>) (Supplemental Text 1). As with all synchronic cross-cultural surveys, each eHRAF culture was assigned an explicit historical contextualization: a single time period directly related to the available documents in the eHRAF cross-cultural files, often referred to as an ethnographic present (Ember and Ember 2009; Swanson 1980). The exploration of the eHRAF cultures resulted in the extraction and analysis of 65,432 paragraphs across the eHRAF cultures under examination (24,763 from imagistic mode cultures and 40,669 from doctrinal mode cultures).

The archaeological aspect of this research centred on the collection, categorization and analysis of material correlate variables from a sample of 50 previously excavated and well documented archaeological site-phases from the Epipalaeolithic to end of the Pottery Neolithic (ca. 20,000-5,300 B.C.) in south-west Asia—encompassing present-day Jordan, Lebanon, Syria, Israel, State of Palestine, and southeast Turkey (Figure 1, Table 1, and Table 3). We used the archaeological material culture to generate a dataset of the instances and patterns of ritual activity, subsistence practice, and social complexity as the agricultural transition progressed in south-west Asia.

For the purpose of uniformly coding, recoding and classifying the ethnographic and archaeological material correlates, categories of absence or presence (0/1) and intensity scales (e.g. 0-3) were used to identify and record the data (Supplemental Text 1). For example, categories of absence or presence (0/1) were used to identify and classify the main subsistence strategy of each culture such as herding (0 = absent, 1 = present) and intensity scales were used to record in more detail aspects of each culture's subsistence strategy—e.g. animal herding intensity (0 to 2) where 0 = no evidence of animal herding: the group are reliant on the hunting of wild animals, 1 = some herding: the recording of herded animals that form a large component of the meat protein intake, along with hunted animals and 2 = intense animal herding: the recording of a substantial presence of herded animals (especially cattle) and little evidence of hunting. In addition, each eHRAF culture and archaeological site-phase were categorized in terms of three groups size measures: (1) less than 150 people, (2) 150 to 500 people; and (3) 500 to 5, 000 people (Dunbar 1993; 1992; Hassan. 1981; Kosse 1989, 1994). In our MCA method, the categorized eHRAF cultures were used to construct a GLM matrix of binary (0, 1) responses representing the known imagistic or doctrinal indicators, which could be used to classify the archaeological site-phases.

### ***Statistical Analysis***

All of the collected data were subject to three complimentary statistical techniques: multidimensional scaling (MDS), principal component analysis (PCA) and generalized linear modelling (GLM). First, MDS was used to provide a general picture of how the eHRAF cultures and the archaeological site-phases, respectively, separated in relation to the recorded ritual, subsistence, and social complexity variables. Secondly, PCA was used to identify the specific sets of ritual, subsistence, and social complexity variables that were responsible for the separation in the eHRAF cultures and the archaeological site-phases. Lastly, GLM was used to examine which set variables (identified via PCA) represents the best predictor of mode of religiosity (in the absence of ritual frequency and arousal data) and to generate the percentage probability of each ethnographic culture or archaeological site-phase reflecting a culture engaged in an imagistic or doctrinal mode of religiosity.

### ***Multidimensional scaling (MDS)***

Multidimensional scaling (MDS) is a dimension reduction method that produces coordinates in dimensional space that best characterize the structure of a dissimilarity matrix, using a Gower similarity coefficient (Baxter 1994; Davidson 1983 Gower 1971a, 1971b). The Gower similarity compares two cases  $i$  and  $j$  and coefficient is defined as:

$$S_{ij} = \frac{\sum_K (W_{ijk} S_{ijk})}{\sum_k (W_{ijk})}$$

where  $S_{ij}$  is the similarity between two individual cases; where  $S_{ijk}$  is the influence of the  $k$ 'th variable,  $w_{ijk}$  is the weight of the  $k$ 'th variable (0 or 1) (Baxter 1994; Gower 1971a, 1971b). In our study, MDS was conducted using the R-statistics package (<https://www.r-project.org/>) (Supplemental Text 2). It was employed to examine grouping patterns relating to each set of ritual activity, subsistence, and social complexity variables from the ethnographic cultures and the south-west Asian archaeological site-phases through the generation of three-dimensional dissimilarity matrices—presenting the multivariate distances of individual cases in relation to the first three principal components.

### ***Principal Component Analysis (PCA)***

Using orthogonal transformation, PCA enabled us to identify the sets of variables that accounted for the maximal amount of variance in the data sets, in terms of a complementary set of score and loadings (Abdi and Williams 2010; Esbensen and Geladi 1987; Jolliffe 2002; Ringner 2008; Saporta and Niang 2009). Through the production of correlation circles and factor maps, we were able explore the relationship of plotted individual eHRAF cultures, as well as archaeological site-phases, to each other in terms of the PCA identified variables. As the recorded ethnographic and archaeological information reflected categorical data (e.g., 0 or 1) and ordinal data (using 0-2 or 0-3 scales), the variables were converted to normal quantile variables to give attractable distributions (normal distributions), enabling standard PCA analysis to be performed. This was done using the *FactoMineR*-package for multivariate analysis (<http://factominer.free.fr/>) (Supplemental Text 2). PCA was carried

out for each set of ritual, subsistence, and social complexity variables for the ethnographic and archaeological databases using mode of religiosity (for ethnographic cultures) and archaeological cultural horizons (e.g. PPNA) as the identifying factors. Central to our MCA approach, PCA enabled us to identify the specific sets of variables that were responsible for the separation in the samples of eHRAF cultures and archaeological site-phases and provided a defined set of variables that could be used in generalized linear model (GLM) matrices to classify the unknown archaeological cases in terms of the two modes.

### ***Generalized Linear Modelling (GLM)***

GLM is a multilevel binary regression statistical technique that provides a model that best accounts for the variance observed in the sample (Agresti 2007; Dobson 2002; Field 2005; Howell 2009; McCullagh and Nelder 1989). GLMs are based on an assumed relationship (link functions) between the mean of the response variable and the linear combination of the explanatory variables (Dobson 2002; Guisan et al. 2002; McCullagh and Nelder 1989). GLM is an extension of the standard Least-Squares Regression; the difference being that Least-Squares Regression assumes that residuals follow a normal (Gaussian) distribution, whereas, GLMs do not assume that residuals follow a normal distribution and can be used to model continuations, ordered and unordered data. Thus, GLMs provide a multivariate statistical method for modelling data that represents a number of probability distributions including: Gaussian, inverse Gaussian, normal binomial, negative binomial, Poisson, and gamma distributions (Baxter 1994; Guisan et al. 2002; Venables and Dichmont 2004). For the purpose of hypothesis testing in our MCA method, GLMs were a particularly applicable formal inferential statistical technique as they offered us a predictor models to analyze archaeological and ethnographic data which (are often) not represented by classical Gaussian distributions. GLMs are fit to data by the method of maximum likelihood, providing the percentage probability that an unknown sample can be classed in terms of a particular known category.

However, GLMs can be subject to overfitting which happens when a model is extremely complex, usually by having too many parameters relative to the number of observations, the result being that the GLM cannot identify the important variables responsible for the separation between cases and

describes random error or “noise” produced by the inclusion of non-significant variables (Bourne et al. 2007; Gusian et al 2002). To limit this potential error, it was necessary first to use PCA to identify the main set(s) of variables responsible for the distinction between cases.

We utilized the specific set of known imagistic and doctrinal cultures as a threshold to make a binary function, divided into imagistic or doctrinal (0 or 1), and a GLM (binary regression) was applied to it. Once the cases had been divided in terms of this binary relationship, the variables identified using PCA were used to apply a multilevel binary regression via the GLM. Firstly, the GLM was applied to the known ethnographic sample to test the appropriateness of each GLM; to examine how successful the model was at correctly categorizing the known imagistic and doctrinal cultures. Secondly, the GLM was applied to the unknown archaeological site-phases. The function *cbind* in the R-statistics package (<https://www.rdocumentation.org/packages/base/versions/3.4.1/topics/cbind>) was used to create a matrix by binding the column vectors containing the binary numbers 0 and 1 assigned to the known ethnographic cases (Supplemental Text 2).

### ***Testing the GLMs via Receiver Operating Characteristic (ROC) curves***

Each GLM was (initially) tested by plotting Receiver Operating Characteristic (ROC) curves (Supplemental Text 2). An ROC curve makes it possible to assess the accuracy of the received predictions by plotting the true positive rate against the false positive rate (Beerenwinkel et al. 2005; Metz 1978) (Figure 2). The standard two-dimensional ROC curve is a graph of the proportion of positive responses in the sample plotted against the proportion of false responses—i.e. false rate. The origin point (0, 0) represents a situation of no positive classifications being assigned; such a classifier commits no false positive errors but also gains no true positives (Beerenwinkel et al. 2005; Bewick 2004; Fawcett 2005; Hartley et al. 2006; Johnson 2004; Metz 1978; Li et al. 2006). In terms of an initial visual inspection, the closer the ROC curve follows the left-hand border and then the top border of the ROC space (making a right angle), the more accurate the test—that is—the more appropriate the model is for generating true predictions (Figure 2). In a perfect test, an ROC curve would start at the origin (0, 0), go vertically up the y-axis to the (0, 1) coordinate and then horizontally across to the (1, 1) coordinate (Beerenwinkel et al. 2005; Bewick 2004; Fawcett 2005; Hartley et al. 2006; Johnson

2004; Metz 1978; Li et al. 2006). The closer the ROC curve tends toward a 45-degree diagonal line, the less accurate the test (Figure 2). A diagonal line (going from 0, 0 to 1, 1) represents the random assigning of classes. For example, if a classifier randomly predicts the positive class half the time, it can be expected to get half the positives and half the negatives correct; this results in the point (0.5, 0.5) in ROC space (Figure 2).

Apart from visually assessing GLMs using ROC curves, a common method to examine the appropriateness of a GLM is to calculate the area under the ROC curve (AUC) (Bradley 1997; Hanley and McNeil 1982; Roomp et al. 2006). The AUC value is always between 0 and 1 (or 0 to 100% of cases classified correctly). The larger or higher percentage (i.e. closer to 1.0 or 100%) the AUC is, the better the GLM's predictor power. The perfect predictor model will result in an AUC of 1 (or 100% of cases classified correctly), while a model producing random classifications will produce an AUC of 0.5 (50% of cases classified correctly) or less. A valid predictor model should have an AUC greater than 0.5. (or 50%); the closer AUC is to 1 (100% of cases classified correctly), the better the predictor model is at classifying the cases correctly (Bradley, 1997; Guo et al. 2006; Hanley and McNeil 1982; Roomp et al. 2006). The combination of recording the true positive versus the false positive rate and calculating the area under the curve made it possible to use ROC curves to examine the validity and performance of the GLMs employed in our research. The *ROC* and *ROCR* libraries in R-statistics package (<https://cran.r-project.org/web/packages/plotROC/index.html>) were used to plot the (ROC) curves, as well as calculate the AUCs and percentage of cases classified correctly.

### ***Testing the GLMs via half normal Q-Q plots***

We implemented an additional level of validation for all GLMs by generating half normal Q-Q plots using the method outline by Collett (2014). A half normal Q-Q plot is a plot in which the residuals are arranged in ascending order and plotted against an approximate of their expected values. A half normal Q-Q plot provides a formal diagnostic assessment of the model's *goodness of fit*. It centres on plotting the ordered absolute values of the Pearson residuals (X-axis) against the corresponding half normal quantiles (Y-axis). A half normal Q-Q plot simulates points in relation to a confidence envelope and a line which shows the means of the simulated values. The confidence envelope is such

that if the fitted model is correct the plotted points are likely to be located within the limits of the confidence envelope; generally a 95% confidence envelope is preferred for testing simulated points.

Using the simulated confidence envelope, the half normal Q-Q plot can be (visually) interpreted without having to make assumptions about the distribution of the residuals. Moreover, the generation of a number of outliers outside the simulated confidence envelope indicates that the fitted model is not appropriate to make reliable predictions/classifications. In addition, the closer the points are to the line showing the means of the simulated values, the more appropriate the model. However, Collett (2014) points out that even with a fit for purpose model, the residuals used in constructing a half normal Q-Q plot may not be approximately normally distributed. Thus, a half normal Q-Q plot of the residuals will not necessarily result in a straight line (an ideal line of the means of the simulated values) for the simulated points.

In our MCA approach, the generation of half normal Q-Q plots (Supplemental Text 2) offered a critical validation of each GLM using the means of Pearson residuals obtained by simulation under the assumption that the model is correct along with a 95% confidence envelope enabling us to (1) assess if the Pearson residuals from the fitted model (simulated points) are within the 95% confidence interval; (2) examine the model residuals in relation to a mean; and (3) identify outliers. For example, the half normal Q-Q plot generated for the GLM used to classify the unknown archaeological cases (Figure 3) shows that all the simulations (the points generated) lie within the 95% confidence envelope (dotted line) and are close to the mean line—with none of the simulated points as outliers. This half normal Q-Q plot demonstrates that the model based is appropriate to classify the unknown archaeological site-phases in terms of the two modes.

## RESULTS

### *General trends from the statistical results: MDS and PCA*

In general, the MDS analysis of the eHRAF data demonstrated a statistical separation between the known imagistic and doctrinal cultures, even in the absence of ritual frequency and arousal data. For example, Figure 4 shows a separation between the imagistic and doctrinal cultures based on sets of ritual variables. In addition, PCA revealed the maximum separation between the imagistic and doctrinal cultures based on five specific variables. Interestingly, three out of the five of the PCA identified variables were subsistence variables—i.e. hunting, agriculture/cultivation (both recorded in terms of: 0 = absent, 1= present) and crop intensity at two (intensive cultivation/agriculture with domesticated staples) (Supplemental Text 1). The imagistic and doctrinal cultures were distinguished (generally) in terms of (1) the imagistic groups that engaged in a hunting and gathering subsistence strategy and practiced secondary mortuary practices (including grave disturbance, excarnation/de-fleshing or re-burial) and (2) the fully sedentary doctrinal groups that engaged in intensive farming (including a range of cultivated crops and herded animals) and provided evidence of a long-term food storage strategy, private food cooking, resource monopolization, communal ritual structures, and cemeteries.

The MDS and PCA analysis of the archaeological dataset resulted in site-phases from different cultural horizons being (commonly) separated in terms of (1) the site-phases that provided evidence of hunting and gathering, individual burials, and flexed burials and (2) the site-phases that provided evidence of intensive agriculture (including a range of cultivated crops and herded animals), communal ritual structures (ritual buildings and/or monuments), and storage of cultural knowledge (deliberate actions to externally store or transmit cultural knowledge, for the purposes of preserving and transmitting) (Supplemental Text 1). These statistical distinctions generally characterize a separation between Epipalaeolithic site-phases and the MPPNB, PPNC, and PN site-phases, respectively.

The variables identified by PCA that were common to the eHRAF and the south-west Asian archaeological dataset resulted in two broad groups: those with evidence of mobile or semi-sedentary hunting-gathering, versus those with evidence of intensive agriculture, a high level of sedentism, and communal ritual structures. The former group consisted of the known imagistic cultures from the eHRAF and the Epipalaeolithic archaeological site-phases. The latter group consisted of the known doctrinal cultures from the eHRAF and the Neolithic archaeological site-phases. We suggest that archaeological material culture evidence of intensive cultivation, fully sedentary groups and communal ritual structures reflect evidence of the emergence of a more doctrinal mode in the sample of site-phases we used. Although expected in relation to the archaeology of the agricultural transition, the identified hunter-gatherer versus agriculturalist divide is interesting as a parallel subsistence divide was also identified as marking a distinction between the known imagistic and doctrinal ethnographic cultures, suggesting a relationship between subsistence strategies (as it would be identified archaeologically) and each of the two modes. In relation to the analysis of the samples based on our three group size categories, relationship between the lowest groups size (less than 150) and the known imagistic cultures, as well as the archaeological site-phases which were classified as imagistic was identified. The eHRAF cultures and archaeological site phases reflecting the largest groups size category (500 to 5, 000) clustered distinctly when examined using the social complexity and ritual variables. As all of the eHRAF cultures and site-phases in this groups size category were classed as doctrinal, we can which suggests a connection between particular types of ritual activity, levels of social complexity and the larger (mainly doctrinal) populations. Our (initial) results indicate that population size—particularly in relation to our smallest and largest group size categories respectively—relates to aspects of ritual/religious practice, subsistence strategy, and levels of social complexity which may be indicative of mode of religiosity in archaeological material.

#### ***General trends from the statistical results: GLM***

The assessment of the GLMs using the ROC curves and half normal Q-Q plots for the eHRAF data showed that the ritual variables recorded (other than ritual frequency and dysphoric arousal level) from the known sample can be used to correctly distinguish between imagistic and doctrinal cultures

with a success rate of 77% of the *known* eHRAF cultures correctly classified. This result is promising as it demonstrates that previously classified modes cultures can potentially be correctly identified in the absence of ritual frequency and dysphoric arousal level information using a recorded set of ritual variables.

For the eHRAF dataset, the GLM results demonstrated that using a set of five variables (identified via PCA) selected from all categories (ritual, subsistence, and social complexity combined) represented the most appropriate manner to distinguish between the imagistic and doctrinal cultures, with 85% (29 cultures) of the known eHRAF cultures correctly classified (Table 2). As previously noted, of the five PCA identified variables used in the GLM, three of the variables were subsistence variables. This result further reinforces the findings from the MDS and PCA that demonstrated that subsistence variables represent a dominant set of distinguishing variables in the context of the known eHRAF cultures. Similarly, the assessment of the GLMs for the south-west Asian archaeological dataset showed that a GLM based on a set of PCA identified subsistence variables represented the best predictor model for classifying the recorded archaeological site-phases in terms of the two modes; with 78% of the known eHRAF cultures correctly classified using the archaeologically identified set of subsistence variables.

#### ***Percentage generated from the GLMs of the archaeological databases***

GLMs based on the PCA identified variables from the archaeological dataset classified all of the 13 Epipalaeolithic site-phases as imagistic, with high percentage probabilities (Figure 5 and Table 3). In addition, 5% of the Neolithic site-phases (2 site-phases) were classified as imagistic, 90% (33 site-phases) as doctrinal and 5% as indeterminate [a percentage probability result of 40 to 59% as the sample cannot be considered positively as one of the two classifier cases (imagistic or doctrinal culture)]. Of the PPNA site-phases, 25% (2 site-phases) were categorized as imagistic and 75% (6 site-phases) were classified as reflecting doctrinal cultures (Figure 5 and Table 3). Of the PPNB site-phases, 90% (18 site-phases) were categorized as doctrinal cultures and 10% (2 site-phases) were classified as indeterminate. Furthermore, all 5 of the Early PPNB site-phases were classified as doctrinal. Of the Middle PPNB site-phases, 90% of the site-phases (9 site-phases) were categorized as

doctrinal and 10% (1 site-phase) were classified as indeterminate (Table 3). Of the Late PPNB site-phases, 75% (4 site-phases) were categorized as doctrinal and 25% (1 site-phase) were classified as indeterminate. All PPNC and PN site-phases were categorized as doctrinal cultures (Figure 5 and Table 3).

For the archaeological dataset, the highest percentage of doctrinal classifications occurs in the PPNC and PN site-phases—both cultural phases with 100% of site-phases classified as doctrinal (Figure 5). However, the high number of doctrinal classifications for the PPNB site-phases (at 90%) is, also, interesting as the PPNB period is connected with a significant population increase, high intensity agriculture, and proto-urbanization (Asouti and Fuller 2012; Bellwood 2005; Fuller et al. 2011; Harris 2002; Kuijt and Goring-Morris 2002; Peters and Schmidt 2004; Rollefson 1998a, 1998b, 1998c, 2000).

The GLM results demonstrated a clear distinction between site-phases with evidence of groups engaged in hunting-gathering versus those with evidence of an intensive agricultural subsistence strategy—with the former (generally) classified as imagistic and the latter (mostly) categorized as doctrinal. The results support the claims and previous findings by Whitehouse (2004) and Atkinson and Whitehouse (2011) in relation to the connection between the engagement in an intensive agriculture subsistence strategy and the doctrinal mode. In this regard, we can suggest that subsistence evidence from the archaeological record can potentially be used to predict mode of religiosity. In relation to the hypothesis central to our research, the percentages generated by the GLMs suggest a general shift from an imagistic to a more doctrinal mode of religiosity associated with the agricultural transition in south-west Asia (Figure 5 and Table 3). From the samples we used and the MCA method we employed, we can assert that (as the agricultural transition progresses) we can observe a decrease in imagistic classifications and an increase in doctrinal classifications.

## **DISCUSSION AND CONCLUSION**

In this paper, we have outlined how the MCA method makes it possible to use archaeological data to extend hypothesis testing beyond the ethnographic or historical record. In doing so, we have shown how MCA can be used effectively in large meta-analytical studies incorporating material evidence from a number of archaeological and ethnographic data sources. We have demonstrated that by, identifying and statistically modelling common aspects of material culture, we can use archaeological material culture to bridge the gap between the theories developed and tested using ethnographic material and actual archaeological material culture—resulting in a truly interdisciplinary method of testing hypotheses. To illustrate our approach we have shown how the MCA method which centres on integrating ethnographic and archaeological data in the context of an analytical framework can be readily employed to examine the ritual and non-ritual archaeological material culture that may indicate the emergence of an increasingly doctrinal religion during the agricultural transition.