**Intellectual merit**

Social interactions are the foundation for the economic, cultural, political, and reproductive relationships that are so vital to human adaptive systems. Archaeology is uniquely equipped to track the development of these interactions across vast expanses of space and time. Guided by the work of other social scientists, archaeologists have come to recognize that if interactions are conceived as formal networks—essentially entities linked together through relationships—a powerful array of methods and concepts from the network sciences and graph theory become available to quantitatively describe and visualize the structure of past social relationships. Here, we outline a Social Network Analysis (SNA) approach to the reconstruction of Upper Paleolithic social landscapes. Our focus is the Magdalenian (ca. 20,000 to 14,000 cal BP) of western and central Europe. This period witnessed both a rapid expansion of human populations from core areas after the Last Glacial Maximum (LGM) and the creation and circulation of an unprecedented abundance and diversity of exotic materials, items of personal ornamentation, and portable decorated objects. The chronological and geographical distribution of these items imply extraordinarily rich intra- and inter-regional social interactions. As a case study for our SNA approach, we concentrate on one class of personal ornamentation that appears across the Magdalenian world: perforated disks. Not only are these forms of visual display relatively common (over 200 whole and fragmentary disks are known thus far), but their distribution, design details, and the materials from which they are made suggest changes over time and space in both the social meaning and patterns of connections among groups.

The social organization of Magdalenian groups has received much attention from archaeologists. Within this context, a nascent community of scholars recognize the value of applying an SNA perspective, rooted in the methods and concepts of network science, to Magdalenian lifeways (and to Paleolithic societies more generally). Importantly, this approach can help compare and evaluate both etic models of Magdalenian social networks, which tend to emphasize reactionary adaptations to harsh, rapidly changing climates and seemingly unfamiliar fauna, flora, and topography, and emic models, which view Magdalenian peoples as active agents who purposely used material culture to negotiate social relationships and celebrate the skills, personal qualities, and knowledge that allowed them to successfully colonize and adapt to new environments.

We seek NSF funds to (1) assemble a database of ~200 digital images of Magdalenian disks; (2) construct an open-access, web-based application that uses image analysis, machine learning, and clustering algorithms to identify stylistic patterns among 2D digital representations of Magdalenian disks; and (3) develop custom plugins within the open-source graph analysis platform Gephi to produce visual representations of, and quantitative descriptors for, networks at multiple scales and according to user-selected attributes that are unique to an archaeological context. Both applications will be freely available to scholars and will support user-specified analyses and the upload and incorporation of additional disk images as they become available. While our focus here is on Magdalenian disks, this procedure can be readily applied to any type of engraved artifact represented by a 2D digital image (e.g., engraved scapulae, stone plaquettes, *contours découpés*, etc.). Concurrent with the development of the image analysis application and the custom-coded plugins, we generate several predictions about network structure and, thus, the role(s) of perforated disks and social networks in the lives of Magdalenian groups. We test these predictions to document how the reactive and proactive responses of Magdalenian peoples to external climatic forces and internal social dynamics influenced network formation and upkeep. The intellectual merit of our study, then, lies in its combined application of (1) the methodological tools of SNA within an Upper Paleolithic context; (2) image analysis and machine learning to artistic motifs; and (3) a dynamic, open-access model of data dissemination.

**Outlining the research problem**

*The role and materiality of social networks among small-scale societies*

We work from Schortman’s (2017: 268) premise that social networks “are the means by which people exercise agency as they cooperate in mobilizing economic, political, and cultural resources in support of shared objectives.” While he employs this conceptualization in the context of state-level societies, we think it applies equally well to foragers and other small-scale societies, whose social networks facilitate solutions to collective problems like resource shortfalls, mate access, information acquisition and dissemination, conflict resolution, and political turmoil (Gamble, 1982; Gould, 1982; Wiessner, 1982; Hamilton et al., 2007; Funk, 2011; Apicella et al., 2012; Salali et al., 2016; Kramer et al., 2017). While these common challenges are strong social adhesives, other factors like competition, disease, and interpersonal conflicts can disrupt social groupings. What is more, there appear to be cognitive, temporal, cultural, and ecological limits to the number of social ties that can be meaningfully cultivated (Chagnon, 1975; Johnson and Earle, 1987; Kosse, 1990; Lovis, 2016). To manage these incredibly complex relationships, people materialize their position and role within a social network through the stylistic manipulation of objects.

Broadly defined, “style” is a visible and variable quality of an object that people create or decorate. More specifically, our study draws on Wiessner (1990: 105) and Bowser (2000: 242), who respectively characterize style as “a means of non-verbal communication” and “a set of rules to be manipulated, and a set of choices to be made,” in the negotiation of one’s social identity. At the same time, we draw on Carr’s (1995) unified theory of artifact design. Ethnographic data show that people regularly convey individual (Gould, 1969; Munn, 1973; Ray, 1977; Tonkinson, 1991; Oakes and Riewe, 1998), group (Weyer, 1932; Massola, 1971; Wiessner, 1989; Davis and Prescott, 1992), and broader cultural identities (Wiessner, 1982; Saitoti, 1988; Rick and Jackson, 1992) through visual languages. Wiessner’s (1982, 1983, 1989) social psychological approach argues that people actively seek to define themselves relative to others. Thus, the choice to use similar or identical materials and/or decorative elements reflects social similarity, while the use of different materials and/or decorative elements indicates social dissimilarity. Bowser’s (2000) ethnographically informed perspective reveals that people consciously (that is, actively) use style to simultaneously convey more than one aspect of their social identity. In the case of the small-scale, segmental Achuar and Quichua societies of the Ecuadorian Amazon with whom she works, women embed into their pottery information about both their ethnic identity and their political allegiance through specific, and distinct, stylistic elements. Members of these closely interacting cultural groups usually understand one another’s messages, including nuanced details that arise when a woman’s ethnic identity is different from her political allegiance. Sinopoli (1991) likewise reports that patterns of stylistic variation differ among individual components of Southern Paiute and Gosiute arrows, which suggests that a single object can communicate social information on more than one level (see also Longacre, 1981; DeBoer, 1990; Chiu, 2007). Carr (1995) argues that the conveyance of social information through visually perceptible material attributes occurs both actively *and* passively through a complex interplay of physical and contextual visibility (cf. Sackett, 1990). Of most relevance here, however, is his insight that the transmission of information through material culture requires that socially important attributes of objects be visible at levels commensurate with their intended audience. Attributes that are highly visible, both physically and culturally, can convey social information to many people, while those that are less visible do so to fewer, and usually more familiar, people.

Several researchers argue that the most conspicuous stylistic differences typically appear along the boundaries of geographically adjacent groups, a finding that supports the view that social differentiation is often strongest between groups that regularly encounter and compete with one another (Hodder, 1977; Wobst, 1977; Sinopoli, 1991). In contrast, the use of very similar styles within groups both reinforces solidarity and helps to avoid internal conflicts and fission. Hodder’s (1979) ethnographic work among cattle herders in the Baringo District of western Kenya reveals the influence of economic competition and population density on the use of visual languages as a form of boundary maintenance. Under conditions of resource abundance and low population density, stylistic differences between groups are relaxed, even at boundaries, because there is little need to reinforce internal cohesion for the sake of resource defense. When resources are limited and population density is high, the maintenance of social boundaries and internal cohesion become critical, and sharp stylistic differences emerge between groups.

Wobst’s (1977) “emitter/receiver” model and Carr’s (1995) unified theory nicely tie together information transmission and multi-level stylistic (or attribute) variation. Because coarse levels of stylistic variation (e.g., form, material, and/or size) can be more readily seen and decoded from afar, they tend to provide information for and about socially distant groups and/or individuals. Variation at intermediate levels (e.g., locations and types of decorations) is typically less visible and, thus, often records information for and about socially closer groups and/or individuals. Fine-grained variations (e.g., design details and methods of execution) are only visible at close range and are strongly influenced by unique traditions of learning, so they most likely encode information for and about socially intimate groups and/or individuals. We interpret perforated disks within this theoretical framework to detect patterns in Magdalenian network connectivity.

*The Magdalenian world*

While several regionally-specific chronologies exist for the Magdalenian, most researchers still recognize three major temporal-cultural periods (summarized in Gravel-Miguel, 2017): Lower (ca. 20,000 to 17,500 cal BP), Middle (ca. 17,500 to 16,500 cal BP), and Upper (ca. 16,500 to 14,000 cal BP). The Magdalenian is characterized by a common suite of material culture that includes (1) blade- and bladelet-based lithic technologies; (2) abundant and diverse osseous artifacts; (3) perforated animal teeth, bones, minerals, and marine and fossil shells; (4) portable decorated objects, including carved and engraved bones and stones; and (5) rock art depictions of geometric designs and naturalistic and schematic animals, humans, and plants (Straus, 2000; Enloe, 2001). At its maximum archaeologically recognized extent, Magdalenian material culture was distributed from Portugal to southern Poland and from the western Mediterranean to the (then-dry) North Sea (Kozlowski, 1989; Straus, 2000; Svoboda, 2000; Álvarez-Fernández, 2009). Southwestern France and northern coastal Spain were core areas that saw consistent human habitation from the Last Glacial Maximum (ca. 27,000 to 21,000 cal BP) through the entire Magdalenian (Bocquet-Appel and Demars, 2000; Gravel-Miguel, 2017), although the intensity of this occupation may have been less than previously thought (Barshay-Szmidt et al., 2016). Other regions were inhabited less intensively, or not at all, during the Lower Magdalenian. Over the course of the Middle Magdalenian, human populations increased dramatically in the Pyrenees and, to a lesser extent, in other higher-elevation and higher-latitude regions (e.g., Switzerland; Leesch et al., 2012). By the Upper Magdalenian, people were entrenched in all regions of western and central Europe (Bocquet-Appel and Demars, 2000). This northeast-trending wave of population expansion over time created unique opportunities for colonizers to employ visual languages to negotiate social relationships both among each other and with densely populated core areas. In newly inhabited areas, for instance, exotic materials, personal adornments, and distinctive decorated objects appear frequently (Schwendler, 2004). In regions with relatively established populations, however, such portable visual displays are rarer and, in places, substituted by, non-portable, landscape-based displays (e.g., rock art). This suggests that pioneer and/or peripheral groups relied heavily on portable decorated objects to broker social exchanges, whereas their counterparts in established population centers did not, perhaps due to differences in social structure (Schwendler, 2012). A similar trend—the widespread use of decorated objects during colonization and their corresponding decline after population establishment—is documented in other archaeological contexts as well (e.g., the Lapita Cultural Complex of Polynesia; Kirch, 1991; Wahome, 1997; Chiu, 2007, 2012; Sheppard, 2011).

*Approaches to the reconstruction of Magdalenian social landscapes*

With some exceptions (Schwendler, 2004, 2012; Fuentes et al., 2019; Langley, 2019), Magdalenian researchers typically offer descriptions of, rather than explanations for, visual displays (e.g., Taborin, 1993; Fritz, 1999a, 1999b; Eriksen, 2002; Vanhaeren et al., 2005; Castelli, 2010; Rivero Vilá, 2015; Rivero Vilá et al., 2019). Others examine object characteristics and production sequences in great detail (Fritz, 1999a, 1999b; Clotte, 2001; Rivero Vilá and Sauvet, 2014; Rivero Vilá, 2016; Sauvet and Rivero Vilá, 2016). However, while spatial patterning in visual displays is widely acknowledged as evidence for social networks (Bahn, 1982; Fritz et al., 2007; Álvarez-Fernández, 2009, 2016; Rivero Vilá and Sauvet, 2014; Gravel-Miguel, 2016), solid actor-based interpretations of the underlying social systems are rarely offered. To do so, we need to understand how social entities interact to form complex structures—a phenomenon made more accessible by a network approach that visualizes and quantitatively describes those interactions. Gravel-Miguel’s (2017) study, which examines equid and cervid representations at sites in the Pyrenees, nicely illustrates the potential of an SNA perspective on the analysis of Magdalenian material culture. Based on patterns of design similarities, she concludes that topography and environment largely dictated the distribution of image types and that people ultimately used social networks as safety nets in the face of climatic and environmental uncertainty. We intend to build on this study’s approach in several ways. First, our focus on perforated disks allows us to expand the analysis to encompass not only the Pyrenees (which, after all, preserves the greatest diversity of Magdalenian visual displays), but this artifact’s entire geographic and temporal distribution in western and central Europe. This spatio-temporal framework permits a multi-scalar exploration of social networks that can treat individual stylistic elements, artifacts, sites, or regions as units of analysis to potentially identify changes in disks’ social meaning over space and time. What is more, disks exhibit more consistent and visually distinguishable attributes than the images studied by Gravel-Miguel (2017). This suggests—á la Carr (1995)—that the visibility and cultural importance of the social information conveyed by disks is commensurate with the geographic scale of our analysis. Second, we harness Late Pleistocene climatic and population data to further evaluate, rather than assume *a priori*, the influence of geography, environmental uncertainty, and population density on network structure. Finally, we explicitly consider the possibility that internal factors like social competition and social knowledge contributed to network structure.

**Conceptual approach, methodology, and hypotheses**

*From artifacts to networks*

While network analysis encompasses a wide variety of approaches, they all assume that entities cannot be fully understood outside the context of their relationship(s) with other entities. A network can thus be conceptualized as a series of nodes (i.e., entities) connected by the ties (i.e., relationships) that bind them together. This node-tie structure expresses not only the presence or absence of ties but their inferred strength and direction at multiple scales. Networks are often visualized and quantitatively described through a graph of points and lines. In SNA, nodes represent social units like individuals or organizations, while ties represent socially significant relationships like kinship or shared identity (Marin and Wellman, 2014). The behavior of nodes is thus dependent on, and influenced by, the flow of social information throughout the network in which they appear. SNA approaches appear in sociological studies of interpersonal relationships as early as the 1930s (Freeman, 2014). While these ideas occasionally crop up in the archaeological literature of the 1960s and 1970s (Irwin, 1977; Irwin-Williams, 1977; Terrell, 1977), applications of formal SNA only gained traction in the first years of the twenty-first century (Brughmans, 2013). Today, SNA is employed across disciplines to tackle a variety of questions including, most recently, the diffusion of information in online social networks (Guille, 2013). This is an exciting time for SNA, as powerful computers and sophisticated statistical algorithms now permit the use of massive datasets to detect previously hidden network structure or, in the realm of online networks, fabricated user accounts (Deng et al., 2016).

To move, as Knappett (2011: 8) puts it, from “network thinking into network analysis,” we must first recognize that formal networks are models—that is, they are representations of real-world phenomena (Brandes et al., 2013: 3). In an archaeological context, the conceptual process of moving from a phenomenon of interest to a meaningful representation of that phenomenon with a network involves (1) the identification of an archaeologically visible behavior; (2) the abstraction of appropriate archaeological data into formal network concepts; and (3) the transformation of those archaeological data into network data (Collar et al., 2015: 6; Mills, 2017: 384; Peeples, 2019: 466-471). Our conceptual approach is outlined in Table 1.

Table 1. Conceptual Approach to Magdalenian Social Networks (see Collar et al. 2015).

|  |  |
| --- | --- |
| Phenomenon | Social interactions and their role in Magdalenian societies |
| Conceptualization | Similarities and differences in the style of disks at coarse-, intermediate-, and fine-grained scales reflect the strength, intensity/duration, and/or type of social contact |
| Data representation | Nodes (individual artifacts, archaeostratigraphic units, archaeological sites, and/or regions) are linked together based on shared stylistic elements |
| Methods/tools | Centrality measures; multi-scalar networks; comparison of network structure with (1) random networks; (2) geographic distance networks; and (3) predicted network structure under documented paleoenvironmental conditions, population densities, and social conditions |
| Temporality | Two time periods: Middle and Upper Magdalenian |

The phenomenon of interest here is the exchange of social information and, ultimately, its role in the lives of Magdalenian peoples. As noted above, we assume that perforated disks acted as vessels for socially significant information and, as such, reflected the transfer not only of items but of the skills, practices, relationships, and meanings embedded within them (cf. DeMarrais, 2004). The movement of social information through material culture thus offers a glimpse of peoples’ decisions as they actively construct a social landscape to suit their needs. The nodes of the network are individual artifacts, archaeostratigraphic units, sites, and/or regions while the ties between them are geographic distance and/or the degree of shared material, form, and design elements. While not all of the disks are associated with high-precision radiocarbon dates, nearly all of them can be securely assigned to either of the ca. 1,000- to 2,000-year time slices of the Middle and Upper Magdalenian. These time periods therefore form the temporal framework for our study.

*Artifact sample*

Table 2 lists the proposed sample of disks. We focus on those artifacts for which high-quality images have been published or for which we can acquire digital images.

Table 2. Proposed Sample of Perforated Disks.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Site** | **Medium** | **# of Disks** | **# of Sides** | **Age** | **Location** |
| Abri Plantade | bone | 1 | 2 | UM | 1 |
| Aurensan | bone | 2 | 4 | MM | N/A |
| Bédeilhac | limestone | 1 | 1 | MM | 2 |
| Chaleux | bone, ivory | 2 | 3 | UM | 3 |
| Chancelade | bone | 1 | 2 | MM | 4 |
| Duruthy | bone | 2 | 2 | UM | N/A |
| Enlène | bone | 55 | 67 | MM | 2, 5, 6 |
| Espalungue/Arudy | bone | 5 | 10 | MM | 2 |
| Gönnersdorf | slate, antler, ivory | 22 | 31 | UM | N/A |
| Gourdan | bone | 3 | 6 | MM | 2 |
| Hollenberg-Höhle 3 | lignite | 1 | 1 | UM | N/A |
| Isturitz | bone, sandstone | 38 | 70 | MM | 2 |
| Isturitz | bone | 3 | 6 | UM | 2 |
| Kesserloch | bone, lignite | 3 | 4 | MM | N/A |
| La Garenne/Saint Marcel | bone | 1 | 1 | MM | N/A |
| La Madeleine | bone | 3 | 3 | MM | N/A |
| La Tuilière | bone | 1 | 2 | MM | 7 |
| La Viña | bone | 1 | 2 | MM | 2 |
| Laugerie-Basse | bone | 6 | 11 | MM | 2 |
| Le Mas d’Azil | bone, jet | 41 | 81 | MM | 2, 8 |
| Le Morin | bone | 1 | 2 | UM | 9 |
| Le Portel | bone | 2 | 4 | MM | N/A |
| Les Combarelles | bone | 1 | 1 | MM | N/A |
| Les Trois Frères | bone | 4 | 4 | MM | 5 |
| Llonín | bone | 1 | 2 | MM | 2 |
| Lortet | bone | 3 | 6 | MM | 2 |
| Lourdes | stone | 1 | 2 | MM | N/A |
| Montastruc | bone | 6 | 9 | MM | N/A |
| Petersfels | lignite, bone, ivory | 6 | 6 | UM | N/A |
| Saint Eulalie | bone | 1 | 1 | MM | N/A |
| Saint-Michel/Arudy | bone | 3 | 6 | MM | 2 |
| Total | N/A | 216 | 352 | N/A | N/A |

Ages: UM = Upper Magdalenian, MM = Middle Magdalenian; Location: 1 = Musée d'histoire naturelle, Montauban; 2 = Musée des Antiquités Nationales; 3 = Royal Belgian Institute of Natural Sciences; 4 = Musée de Périgueux; 5 = Musée Bégouën; 6 = Musée de l’Homme; 7 = British Museum; 8 = Musée de Mas d’Azil; 9 = Musée d’Aquitaine.

*Analysis of stylistic elements*

Perforated disks, like many items of Magdalenian personal ornamentation, are characterized by a wide variety of attributes (*sensu* Carr, 1995) or media and design elements (Barandiarán, 1968; Sieveking, 1971; de las Heras et al., 2007; Corchón Rodríguez and Rivero Vilá, 2008). There are nevertheless spatio-temporal patterns in size, medium (bone, ivory, antler, stone), design location (one or both faces; edge versus center), and design type (notches, lines, perforations, figures, geometric motifs) (Schwendler, 2005). We retain these categorical variables of medium, number of decorated faces, and the presence/absence of perforations and couple them with an image analysis approach that permits the extraction and analysis of the engraved designs, which can themselves be further segregated into their constituent units (e.g., edge or center decoration). These variables (medium, decoration face, perforation, digitized engraved designs), taken together as a coherent unit or separately as constituent parts, then serve as input features for SNA. Because differences in the size, medium, and presence/absence of decorations (on one or both faces) and/or perforations can be detected at the greatest distance, we consider these to be “coarse,” or highly physically and contextually visible, design elements that are most likely to communicate information for and about socially distant groups and/or individuals. The location (e.g., edge, center) and type of decorations (e.g., notch, line, circle, figure) can only be identified at closer distances. We therefore categorize these as “intermediate,” or moderately visible, design elements and, thus, those most likely to communicate information for and about socially closer groups and/or individuals. The nuanced stylistic details of the edge or center engravings and the specific type of central figure can be distinguished only at very close proximity. These we classify as “fine,” or low visibility, design elements that most likely communicate information for and about socially intimate groups, individuals, and/or artistic traditions.

To extract the engraved designs, each artifact must first be converted into a 2D digital image, which can be accomplished through digital scans of published images or digital photography of the artifacts themselves. Many high quality disk images are available in research articles and compilations like André Leroi-Gourhan’s (1967) classic, *Treasures of Prehistoric Art*. Those disks without high-quality published images will be digitally photographed at their repositories by graduate students at European institutions. The analysis of digitized prehistoric art is now relatively common (e.g., Domingo et al., 2013; Carrero-Pazos et al., 2016) and is rapidly being incorporated into SNA (e.g., 3D scans of sword hilts; Golubiewski-Davis, 2018). The conversion of the disk images into a format amenable to anthropogenic feature extraction and similarity analysis presents several challenges, including (1) the complexity of the engravings, which include both linear and curvilinear structures and the junctions between them; (2) illumination heterogeneity due to the direction of the light source and/or artifact surface irregularities; (3) incompleteness of the design because of breakage; (4) variation in artifact shape; and (5) non-anthropogenic modifications (e.g., post-depositional scratches and cracks). The best way to address these issues is through a cascaded image processing procedure that involves image enhancement, scale invariant structure extraction and, finally, feature representation. Our preliminary attempts at image enhancement indicate that non-anthropogenic and anthropogenic modifications can be differentiated. We find, for example, that non-anthropogenic modifications appear as filament-like structures that can be captured and suppressed with a Hessian matrix, which classifies pixels into two groups (filament-like or not) through the definition of elongated ellipses based on local structures in the image. The anthropogenic modifications, which appear as ridge-like structures, can be detected and enhanced through the use of a linear combination of steerable image filters. Figure 1 shows an initial attempt at image enhancement for a disk engraved with the image of a chamois from the site of Laugerie-Basse (Graziosi, 1960) in addition to a sample of non-enhanced disk images. Note that the non-anthropogenic background “noise” is suppressed in the Laugerie-Basse image while global, anthropogenic structure is retained and enhanced.

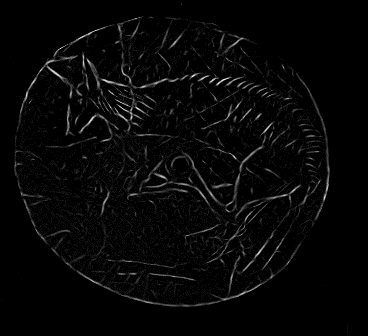
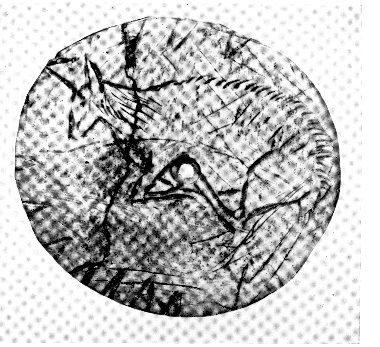
 

Fig. 1. Example of an original scanned image (left) and edge enhanced image (center) of a perforated disk from Laugerie-Basse, and examples of disk styles (right) from Schwendler (2005).

After image enhancement, structure extraction can more precisely detect the centerlines of the anthropogenic modifications at the pixel level in order to represent their form more compactly. To do so, the image is converted into a distance map that represents the relative distance of each pixel from its closest background pixel. The centerline pixels can then be determined by taking local maxima from the distance map. A double threshold is then applied to further prune unwanted line segments and/or isolated point sets (i.e., any remaining non-anthropogenic modifications). The lower-bound threshold identifies and excludes “weak” features that are not connected to the structure of interest—the anthropogenic engraving in this case—while the upper-bound threshold identifies and includes those features that are connected to the structure of interest. This protocol thus retains only those pixels that reflect the anthropogenic structure of the images. Importantly, this methodology employs a multi-scalar approach based on scale-space theory, which explicitly and simultaneously considers images at multiple scales (Lindeberg, 2008). This way, we can focus on fine details such as ridge points at the microscale and global shape at the macroscale. The resulting enhanced images can then be compared for similarity and incorporation into the SNA.

The last step in this process involves the construction of a regression-based machine learning model (e.g., Support Vector Machine or Convolutional Neural Network) that can associate unenhanced anthropogenic engravings in the raw digital images with the “clean,” enhanced anthropogenic engravings in the corresponding processed images. In this case, a supervised learning approach to model training is used, where the algorithms are presented with both the unprocessed and processed images. We will validate the image enhancement process by comparing the results with (1) manually traced contours that accompany published images of disks; (2) digital photographs of disks for which high-quality published images also exist; and (3) scanned images of disks that appear in multiple publications under different lighting conditions. Once the model is trained with our sample of disks, it becomes part of the online application so that users can forgo the time-consuming and computationally intensive process described above and simply convert their images for comparison and SNA with an existing database of disk images.

*The identification of social ties*

As noted above, we assume that similarities in the attributes of disks between nodes reflect the strength (level of social affinity), intensity (frequency or duration of social encounters), and/or type (communication with a socially intimate or socially distant unit) of Magdalenian social interactions. Assessing similarity among nodes for medium and decoration location is based on clearly defined categories. We can easily compare nodes based on the presence/absence and/or frequency (raw or relative) of disks made from particular raw materials and with decorations on one or both faces. Comparisons among nodes based on decorations, however, first require the use of statistical classification. The classes in this case are clusters of shared engraving patterns—motifs—that represent shared social information. Because the way(s) Magdalenian artisans assigned meaning to their visual displays is unknown, the number and type of motifs cannot be predefined. We therefore utilize unsupervised classification, which makes no prior assumptions about class membership. There are two general approaches to unsupervised image classification, dimension reduction and clustering (Olaode et al., 2014). Most appropriate for the complex and multidimensional (potentially over a thousand) data of the disk images are non-linear methods of dimension reduction like Kernel Principal Components Analysis and Multidimensional Scaling, both of which use nearest neighbor algorithms to identify underlying structure. With this approach we can extract the decorations of an entire disk face as a unit of analysis or, alternatively, focus on specific elements like edge decorations and central decorations as separate units of analysis. The associated dimension or component loading scores (either for entire disk face decorations or specific decorative element on a disk face) then serve as attributes for similarity analysis.

Clustering techniques seek to define natural groupings such that images within the same group share more attributes with each other than they do with images from other groups. The stylistic features used for clustering are often predefined by the analyst (e.g., Rivero Vilá and Sauvet, 2014; Gravel-Miguel, 2016). Because we do not know what stylistic features were important to Magdalenian craftspeople, we permit the algorithm to explore the structure of the data without user input. Our analysis focuses on the *k*-means clustering method, which possesses an adjustable parameter *k* that sets beforehand the number of classes (i.e., groups of disks) to be defined. Within the context of Magdalenian perforated disks, this enables “stylistic granularity”—the freedom to determine how much stylistic detail to use in the definition of disk groupings. If a high degree of granularity is desired (e.g., to group together those disks that match precisely across many stylistic details, perhaps even down to the level of individual craftspeople), the algorithm digs deep into the data structure to pull out a larger number of groups, each of which contains fewer disks. Coarse granularity, on the other hand, prompts the algorithm to search for more superficial similarities and, thus, to identify fewer, but more populous, groups of disks. This means that stylistic granularity can be manipulated by changing *k* to suit the needs of the analyst: fine stylistic granularity can be attained by setting *k* ~ *n*, intermediate granularity by setting *k* < *n*, and coarse granularity by setting *k* << *n*. There are several ways to measure similarity with presence/absence, frequency, and continuous data, including Euclidean Distance, the Brainard-Robinson or Jaccard Indices, and Proximal Point Analysis (Östborn and Gerding, 2014; Habiba et al., 2018). It is also possible to produce both unweighted networks (i.e., only the presence or absence of ties based on some predefined similarity threshold) or weighted networks (i.e., similarity defines the strength, or weight, of ties) with these metrics.

*Sample size and the problem of class imbalance*

We must consider, too, the fact that disks are very frequent at a handful of sites but only sporadically represented at many others. While this is as much a behavioral pattern to be explained as a statistical problem to be resolved, the severity of the “class imbalance” issue also depends on the nature of the analysis. One option is to ensure that the sample size at each node is equivalent (or more nearly so) by scaling down to individual disks or up to larger spatial units. Stylistic comparisons between individual disks, for instance, will set the sample size at all nodes to one. Similarity metrics in this case will be based on disk-to-disk stylistic resemblances in face, edge, and/or central decorations (estimated by dimension or component loading scores) when *k* = *n*. Sample size imbalances will also be less pronounced when network nodes are represented by regions (e.g., the Pyrenees, southwestern France, and central Europe) that pool disks from several sites. This still leaves the analytically intermediate—and arguably most behaviorally interesting—sites and archaeostratigraphic units. While the class imbalance issue cannot be eliminated at this scale of analysis, its effects can be assessed through estimates of sampling error. To do so, we can compare similarity metrics derived from two nodes of analytical interest to those from a set of randomly sampled node pairs (say, 100 or 1,000). If the original similarity metric does not deviate significantly from what would be expected by chance, it is more likely to be the result of sampling error (see Peeples, 2011 for an application using the Brainerd-Robinson Index). (This approach can, in fact, be applied at any scale of analysis.) Eventually, the disk data can be supplemented with those of other artifact types once the image analysis protocol outlined here is fully developed.

*Network analysis*

Gjesfjeld (2015) identifies two hurdles to the reconstruction of small-scale, prehistoric forager networks: aggregation and fragmentation. The assemblages of disks at Magdalenian sites are undoubtedly aggregations of multiple episodes and many years of use and discard. To alleviate the effects of aggregation in the absence of unique and precise radiocarbon dates for each artifact, we split the disk sample into two meaningful time periods, the Middle and Upper Magdalenian. This framework provides the best achievable compromise between temporal precision and the maintenance of meaningful sample sizes. While we recognize that these time slices do not reflect the day-to-day social interactions of Magdalenian peoples, we do assume that they represent, and that we can access, the most routinized and consistently reproduced social relationships. Network fragmentation—that is, the loss of nodes due, in this case, to the vagaries of the archaeological record—effectively guarantees that we cannot “see” every relationship in a Magdalenian social network. We address this issue in two ways. First, the boundaries of Magdalenian networks are defined as broadly as possible (i.e., across virtually the entire geographic and temporal range of disks in western and central Europe) to maximize the number of sampled nodes. Second, we evaluate the stability of Magdalenian networks through resampling. If network structure remains relatively stable even in the absence of randomly removed nodes, it is more likely that the preserved nodes reflect the original, prehistoric network with some degree of fidelity.

Network structure can be characterized both quantitatively and visually. Many tools are available to quantitatively describe the structure of networks, the most popular of which are measures of centrality. Common metrics like degree, closeness, betweenness, eigenvector, and random walk centrality permit the identification of nodes that, because they are connected either directly or indirectly to many other nodes, tend to exert a strong influence on the structure of a network. In an archaeological SNA context, such nodes are often interpreted to have access to more social information or prestige, facilitate the transfer of social information, and/or wield influence or power over other nodes (Brughmans, 2010; Fitzhugh et al., 2011; Knappett, 2011; Coward, 2013; Phillips and Gjesfjeld, 2013; Rivers et al., 2013; Mills et al., 2015). Here, central nodes are those artifacts, archaeostratigraphic units, sites, and/or regions that share patterns of disk design with many other nodes. Importantly, centrality measures can be calculated at multiple scales to identify not only the importance of single nodes within a region but groups/cliques of nodes between regions. Once such groups are identified, we can employ less commonly used network metrics like social circle intersects to gauge the influence of groups on one another. Statistical comparisons of network structure take one of two forms: network-wide structural comparisons (e.g., D-measure and graph Laplacian, graph edit, and Hamming distances; Schieber et al., 2017) and comparisons of the distributions of individual, node-level metrics (e.g., centrality, similarity). The visualization of networks as a graph of points (nodes) and lines (relationships) in geographic and social space offers an intuitive way to reveal patterns that are not readily apparent through quantitative descriptions alone. The symbology (e.g., color, size, length) of nodes and lines can also be modified to highlight particular network attributes. There are numerous hypotheses about Magdalenian disk-based network structure that can be explored with these tools. Here, we examine five possibilities, namely that disk networks were: (1) random; (2) driven by geography and topography; (3) established to foster permeable social boundaries; (4) established to enforce and maintain strict social boundaries; and (5) elements of social competition and differentiation.

Hypothesis 1: Magdalenian network structure was random.

This is a set of null models at all scales of analysis (individual artifact, archaeostratigraphic unit, site, and region) and using all elements of disk form and design (coarse-, intermediate-, and fine-grained elements) and is based on randomly generated samples of weighted networks of the same size (number of nodes) and density (number of ties) as the observed weighted networks.

Hypothesis 2: Magdalenian network structure was driven by geographic/topographic proximity.

Distance, rugged terrain, and physical barriers all likely play some role in structuring forager social networks. We expect such “geographical friction,” to use Coward’s (2013: 257) term, to simultaneously decrease the likelihood of establishing social connections and increase the cost of maintaining them (Fitzhugh et al., 2011). If so, this model predicts that similarity metrics at all levels of disk form and design and at all scales of analysis should be negatively correlated with straight line distance and/or the cost of pedestrian travel (estimated by cost surface analysis) between sites or regions (Coward, 2013; Gravel-Miguel, 2016). Deviations from this expectation by one or more design elements and/or at one or more scale of analysis suggest that certain types of social information travelled considerable distances and, perhaps, across substantial geographic and socio-political barriers. Straight line distance between sites is readily available from geographic coordinates, and cost surfaces can be estimated through digital elevation models (cf. Egeland et al., 2010) of modern western and central European landscapes available through the European Environment Agency (2017).

Hypothesis 3: Magdalenian network structure fostered permeable social boundaries.

Foragers, particularly those at higher latitudes, may need to establish and maintain permeable social boundaries to facilitate access to resources, information, and mates (e.g., Pearce, 2014). If Magdalenian networks functioned in this way (that is, as “safety nets”), we expect ties based on coarse and intermediate design elements at all scales of analysis to be frequent and strong between different ecological zones and correspondingly rare and weak between similar or identical ecological zones. This pattern would arise if people, by signaling membership in a wider network of variably related, but periodically cooperative, groups, used disks to access resources that were not available (or available less frequently) in their own territory. A related prediction involves pioneer and core populations. If colonizers maintained close relationships with core areas to alleviate the risks of confronting unfamiliar fauna, flora, and topography, we expect strong ties between fine and intermediate design elements. This would emerge as people sought to maintain a shared identity with the socially close and socially intimate groups of their homeland. Data to define ecological zones are available from PaleoView, which provides estimates of annual temperature (maximum, minimum, and mean), annual precipitation (maximum, minimum, and mean), net primary productivity, and seasonality (maximum temperature/precipitation – minimum temperature/precipitation) for 10-year intervals at a spatial resolution of 2.5° x 2.5° for the last 21,000 years (Fordham et al., 2017).

Hypothesis 4. Magdalenian network structure enforced strict social boundaries.

Foragers are known to physically repulse trespassers and/or regulate access to group membership as a means of territorial defense (Layton, 1986). The incentive to defend territories through strict boundary maintenance may be particularly high under conditions of severe resource stress and/or high population density. If Magdalenian peoples used disks to encourage within-group cohesion and to monitor and enforce social boundaries, we expect to see patches of densely connected, geographically proximate sites with few or no ties between patches of sites. Internal cohesion and boundary maintenance would encourage (or perhaps impel) stylistic conformity among socially close and socially intimate individuals, so ties based on fine and intermediate design elements should be especially strong (although coarse design elements may be shared within groups as well). The PaleoView dataset can be used to create a gradient of resource stress: areas with warmer temperatures, more precipitation, higher primary productivity, and less seasonality are considered more hospitable than those that are colder, drier, less productive, and more seasonal. Using the “rcarbon” package in the R statistical environment (Bevan and Crema, 2018), we approximate population densities with summed probability distributions (Contreras and Meadows, 2014; Crema et al., 2017) of ~900 radiocarbon dates from western and central European Magdalenian-aged sites available through the Canadian Archaeological Radiocarbon Database (Martindale et al., 2019).

Hypothesis 5. Magdalenian network structure reflected social competition and differentiation.

The entry by Magdalenian pioneers into previously uninhabited areas no doubt posed unique challenges. The ethnographic record tells us that people often value, and compete to display, the skills, qualities, and information that are most important during challenging times (e.g., those with hunting prowess, knowledge of distant places, or the most secure social ties; Turner, 1894; Wiessner, 1982; Gould and Saggers, 1985; Karklins, 1992; Lourandos, 1997). When performed outside the context of one’s homeland, this competition may result in the reinterpretation of old styles or the invention of new ones as people assert their creativity and negotiate new social identities and allegiances (cf. Chiu, 2012). If such processes were at work among Magdalenian groups, we predict that ties among coarse design elements will be weak within and between peripheral regions as new, and highly visible, styles appear to signal social distinctions among bands. At the same time, vestiges of a shared common culture may be reflected in wide-ranging ties based on the fine design elements that pioneer artisans inherited from core populations.

**Image analysis application and SNA plugins**

*Image analysis application*

This web application houses our raw-image-to-processed-image predictive algorithm and will be made freely available online to all registered researchers. It will be capable of (1) accepting raw artifact images in a variety of formats (JPEG, TIFF, BMP) for processing into “clean,” compact representations ready for similarity analysis; (2) comparing uploaded images to our database of 200+ disk representations; and (3) conducting similarity analysis. For the latter procedure, users can elect to employ dimension reduction, which results in continuous variables (i.e., loading scores) for use in similarity measures, and/or *k*-means clustering, which produces categorical motif groups whose frequency can be compared among nodes. If the latter option is chosen, the user then has the freedom to determine the level of stylistic granularity through the manipulation of the *k* parameter. All of the resulting data can be downloaded in CSV format for use in Gephi. The web application will be built with Django, which is a Python Web framework. The image processing on the backend will also be implemented in Python to allow easy integration with the “LIBSVM” support vector machine library (Chang and Lin, 2011) and the Convolutional Neural Networks library in TensorFlow (Abadi et al., 2016). The application and the data will be stored on the webserver hosted by The Art of Security, Privacy, and Networking (SPartAN) lab at UNCG. The digital images will be archived and made accessible to the research community and interested public parties in the Digital Archaeological Record (tDAR) in a customized, searchable database.

*SNA plugins*

Gephi is an open-source web platform that permits the visualization and analysis of graphs. While Gephi is powerful and dynamic, it lacks several features that are required for our study. Thankfully, there is an active community of developers, and the platform itself supports a “repository of code, collection of builds, and a library of API references all designed to help…extend Gephi’s functionalities.” We will therefore develop several plugins. The first will permit node viewing and analysis in strictly geographic space. An existing plugin, GeoLayout, will be modified to support the visualization and SNA of several GIS-type layers such as standard maps, paleoclimate raster grids, digital elevation models, and kernel density distributions. The second plugin is closely related and deals with data filtering. Gephi currently allows users to filter categorical data (e.g., regions like Cantabria or Aquitaine or dates like Middle or Upper Magdalenian in our study). Our plugin will permit the creation of spatial filters where users generate one or more areas of interest with customizable polygons. Finally, there are network metrics, including several network-wide structural comparisons, that are not currently supported by Gephi but can be added with additional plugins.

**Results from prior NSF support**

Deng (with G. Amariucai of Kansas State University and S. Wei of Louisiana State University) received NSF support through the Communication and Information Foundation Program for “Harvesting Network Randomness and Diversity” (NSF CCF-1320428, $155,817, 6/1/13-5/31/17).

*Intellectual merit*

This project focused on communication networks, which are naturally dynamic, inherently redundant, and largely unpredictable. While the former two features have long been recognized as a valuable resource for integrity, efficiency, and confidentiality, the latter is often regarded as an incommodity. This project demonstrated that network randomness can be harvested and, together with diversity, exploited to enhance communication security.

*Broader impacts*

The development through this project of a deeper appreciation for the statistical nature of networks can be applied to a broad range of information-assurance objectives, including multipath diversity, network tomography, secure network coding, protocol coding, anonymous routing, and the spread of epidemics. Three MSc students and three REU students, including two women, were supported by this project. Several publications and conference presentations resulted from this project as well (Moharrer et al., 2015, 2016; Khalili-Shoja et al., 2016; Joyce and Deng, 2017; Man et al., 2018), including one (Deng et al., 2015) that was recognized as the best paper at the 3rd International Conference on Building and Exploring Web Based Environments.

**Project timeline and investigator roles**

We seek funding for a total of 36 months with work to commence in January of 2021. The first year will concentrate on the acquisition of disk images, while years two and three will focus on the development of the image processing procedure, the SNA plugins, and the web interface. Table 3 outlines investigator roles and responsibilities.

Table 3. Investigator Roles and Responsibilities.

|  |  |
| --- | --- |
| Task(s) | Investigator(s) |
| Oversee project progress; coordinate team meetings | Egeland |
| Work with UNCG anthropology undergraduate student, the Digital Collections staff at the University Libraries at UNCG, and international graduate students to produce and quality-check the digital scans and photographs. | Egeland, Schwendler, Kim |
| Supervise computer science graduate student in the development of image processing and machine learning procedures | Kim |
| Supervise computer science graduate student in the development of SNA plugins and web interface; run SNA | Deng |
| Collect and process GIS, paleoclimate, and radiocarbon data and integrate them into the SNA platform | Nicholson, Deng |
| Interpret and analyze results | Egeland, Schwendler, Kim, Deng, Nicholson |

**Broader impacts**

*Development of a diverse, globally competitive STEM workforce*

This project involves the training of two MSc students in computer science. The first will work with Kim to develop a novel approach to image analysis and classification, construct and train complex multidimensional predictive models, and build an online platform that incorporates several types of software. For this position, we will seek a student with a background in computer science and strong web and Python coding skills. The other student will work with Deng to conduct SNA with custom-coded plugins on a unique sample of archaeological data. This student will also be enrolled in the Department of Computer Science, as UNCG does not have a graduate program in anthropology. However, the qualifications for this position will include not only solid coding skills and experience in math (especially graph theory) but a degree in anthropology, archaeology, or a related social science. Anthropology programs with strong traditions of melding social science with network science are particularly attractive institutions for potential candidates. Egeland will also serve on the graduate committees of both MSc students. The tasks that these students will engage in, especially machine learning and software integration, are highly valued in today’s tech-heavy job markets. Just as important is the meaningful exposure of a computer science student to the questions of social science and the training of an anthropology/archaeology student in the nuts-and-bolts of computer programming, both of which help create tangible links between the “how” and “why” of software design and use (Havill, 2019; Irving and Askell, 2019). Two of the Principal Investigators (Schwendler and Kim) are women, and they will play integral roles in data collection and analysis. Schwendler is both a mentor for several Master’s-level women at PaleoWest and an avid public speaker. She feels strongly that this project and her role as one of the highest-level women at PaleoWest can be used as a platform to encourage women to pursue research opportunities in archaeology. The computer science program at UNCG has successfully recruited female graduate students for several years (~50% of current MSc students are female), and UNCG Human Resources enlists the support of recruiting services to ensure that the listings for both MSc positions will reach a diverse population of potential applicants.

*Improved STEM education and educator development*

While a spatial dimension is often incorporated into SNA, time (particularly the deep time of archaeology) is less commonly considered, especially by practitioners in computer science. This proposal will in fact change the way Deng teaches a recently developed computer science course (CSC 407/607 Network Analysis). He intends to incorporate the techniques, perspectives, and datasets from this interdisciplinary study into a set of class “mini projects” that will confront computer science students with novel problems related to the social sciences.

*Enhanced infrastructure for research and education*

Given the scope of our analysis (~200 artifacts from ~30 sites), we do not expect this to be the final word on Magdalenian social networks. It is important to note, however, that slight modifications to the image analysis algorithms can accommodate any type of engraved artifact that can be represented as a 2D digital image. We therefore see this project as a step toward a long-term investigation of social landscapes that, through the development of new (image analysis application) and enhanced (Gephi plugins) infrastructure, can eventually extend to other types of Magdalenian artifacts and even into other archaeological times and places.