The Threeway Approximate Spatiotemporal Symmetry (TASS) algorithm: A method for trivariate time series segmentation

Gustav R. Sjobeck

M.A., University of Virginia, 2020

B.S., University of La Verne, 2016

A Dissertation Presented to the

Graduate Faculty of the University of Virginia in

Candidacy for the Degree of Doctor of Philosophy

Department of Psychology

University of Virginia

April 27, 2022

Committee Members:

Steven Boker, Ph.D. (Chair)

Teague Henry, Ph.D.

Adrienne Wood, Ph.D.

James Morris, Ph.D.

Tanya Evans, Ph.D.

Contents

Dedication

This dissertation is dedicated to the memory of my grandparents, Gustave and Edna Sjobeck. I wish I could have called to tell you the news. I hope this dedication reaches you, wherever you are.

Acknowledgements

I cannot adequately express how grateful I am to everyone who supported me through the process of receiving my PhD. First, I would like to thank my advisor, Steve Boker, for guiding me in developing the start of a research paradigm that I will carry with me the rest of my career. With your kind advice and suggestions, you helped sculpt the way I think about psychology research and statistical methods, symmetry and interpersonal behavior, and our responsibility as Quantitative Psychologists. I am so grateful for your unique perspective and I hope that we remain collaborators for the rest of time.

Next, I'd like to thank Teague Henry for helping me figure out how to take what I've learned as a graduate student and use it to start a career in academia. You're a wonderful addition to this department and I look forward to all the work we have planned. I'd also like to thank the rest of my dissertation committee—Adrienne Wood, Jamie Morris, and Tanya Evans. Thank you for taking the time to consider my research ideas and helping me apply them to important real-world problems.

Thank you to the "Quant Kids": Bobby Moulder, Jessica Mazen, Tara Valladares, Christof Fehrman, Elena Martynova, and Laura Jamison. Together we navigated the theoretical and practical headaches that come with statistics and administration. I would not have made it through without our collective emotional support and shared knowledge.

Lastly, I would like to thank my friends and family for so many years of love and support: to my friends, both those who were here in person and those who live in far and distant lands, who made me laugh and kept me from losing my sanity; to my family, who never doubted my abilities and patiently nodded along to long-winded explanations of my research; to my parents, Ed and Danielle Sjobeck, and my sister, Madison Sjobeck, who with loving hands crafted me into the person I am; to my undergraduate advisor and academic father, Glenn Gamst, for making it possible for me to go to grad school in the first place; and to my loving partner, Lauren Ballard, who helped make sure this dissertation was finished in a timely manner. Without all of you, the world would be a different place.

The Threeway Approximate Spatiotemporal Symmetry (TASS) algorithm: A method for trivariate time series segmentation

Introduction

Often the first step taken by researchers interested in dynamic phenomena is to understand where a complex system exhibits invariance. Here invariance refers not only to systems with zero variability but also systems with variability below a given threshold. In this way invariance becomes a tool with very general applicability. One may use invariance to describe even the simplest of analyses in statistics, the independent samples t-test, where the distinction between groups reveals a potential two-group structure where members of one group may be considered similar to each other but different from members of the other group. In the same way, one may use invariance to understand how a structural model with manifest and latent variables may be structured the same across groups. This is often referred to as measurement invariance, especially in the context of constructing psychological tests. Thus, invariance allows for an understanding of structure among elements of a system that interact.

The principles of invariance stated as such are directly applicable to the study of interpersonal relationships. We might say that two people engaged in conversation exhibit moments of invariance, as when one conversant smiles and the other smiles back. Likewise, moments of invariance may appear in intrapersonal relationships, as with the similarity or redundancy in signals from different brain regions. The current study will consider the invariance in signals of both an interpersonal and intrapersonal nature.

The efforts of the current study are focused on a particular kind of invariance: symmetry. Symmetry is one of many ways that the association between signals may be conceptualized. Historically, symmetry has referred to mirroring patterns of behavior. However, symmetry is a concept used in mathematics to encapsulate any kind of invariance, and not just perfect mirror symmetry. For this reason, we will use symmetry broadly to refer to any association between signals, regardless of its particular pattern.

Methods have been developed to determine the extent to which two signals exhibit invariance. The current approach is offered as a general tool for capturing symmetry between multiple time series. Assumptions are made about features of the time series, including the way in which they were collected, but no assumptions are made about the content of the time series. In this way, the approach could be used to answer a large variety of psychological questions, in areas as diverse as social psychology and neuropsychology. Claims can be made about interpersonal phenomena, in the way people interact with others, or about intrapersonal phenomena, in the way some aspects of an individual interact with other aspects of the individual. For this reason, the general term "agent" will be used to refer to the most specific sources which generated the time series. In capturing the symmetry between brain regions, for instance, each region would be considered an agent, since each region supplies a time series.

Previous algorithms emphasize the magnitude of association, and aggregate across all time to get an average amount of association between signals. The current approach, the Threeway Approximate Spatiotemporal Symmetry (TASS) algorithm, emphasizes the segmented nature of symmetry. This means that symmetry is not expected to occur at all times, but to be intermittent. Likely in this case symmetry exists in a segment, where psychological agents are symmetric for several successive moments. The segment forms out of a need for communicative invariance, and breaks according to changes in context. The TASS algorithm allows for symmetry indication between three time series, with a structure that may be extended to more time series. In this way, claims can be made about group symmetry dynamics.

Symmetry as a Kind of Interpersonal Alignment

Researchers studying interpersonal interactions have amassed a literature of findings on the ways in which people interact. For one, conversation partners tend to mirror each other's postures and gestures, a phenomenon often called the chameleon effect (Chartrand

& Bargh, 1999). Recent research by Butler and Randall (2013) suggests the importance of emotion coregulation in the interpersonal relationships of, for instance, children and their caregivers and adults and their romantic partners.

The language which has historically described similarity between psychological agents is rich in parallels to definitions and phrases used in the common vernacular. It is common in daily life to use the phrase "in-sync", for instance, to refer to instances where two people or two electronic devices coordinate their behavior and act similarly. Likewise, swimmers who swim as a coordinated group are referred to as "synchronized swimmers". In this way, synchrony has become synonymous with similarity and coordination, and is often used in psychological research as a broad term to refer to all forms of interpersonal alignment. Synchrony has been found to have prosocial benefits. Rennung and Göritz (2016), in a comprehensive meta-analysis, found a relationship between synchrony and prosocial benefits, which was moderated by intentionality.

Mathematical models of synchrony, however, like those discussed in Strogatz (2003), define synchrony by its periodic nature. Consider perhaps the most edifying example of synchrony: the flashing pattern of fireflies. Here, the behavior is cyclical and rhythmic, and the goal of each member is to coordinate its flashing to the flashing of its neighbors. Some interpersonal behaviors are periodic, but this is just one component to the cornucopia of association patterns that may exist between signals. Instead, we will use the term symmetry to refer to the absolute association between signals.

Symmetry is a phenomenon recognizable by humans. Research in human cognition suggests that symmetry in objects improves the perception speed of human observers (Barlow & Reeves, 1979; Bertamini, Friedenberg, & Kubovy, 1997; Bornstein, Ferdinandsen, & Gross, 1981). Symmetry may exist in many different kinds of human signals, including neural (Keller, Novembre, & Hove, 2014; Koban, Ramamoorthy, & Konvalinka, 2019), physiological (Butler, 2011), and behavioral (Blairy, Herrera, & Hess, 1999; Boker et al., 2009; Boker & Rotondo, 2002) signals. Symmetry could even be

expected between latent variables like daily reported affect and self-esteem.

The association between time series, which has sometimes been called synchrony and other times symmetry, has been measured using a variety of methods. Perhaps the most common is cross-correlation (Smith, 1997), a linear measure of association between two time series which assumes *stationarity*, i.e., that the parameters which describe time series be stable for all intervals of any length. The windowed cross-correlation (WCC) approach, as proposed by Boker, Rotondo, Xu, and King (2002), emphasizes short-term dynamics by assuming *local stationarity*. The result is a matrix of correlations which capture the association between local regions of each time series at various lags. Several metrics which aggregate the correlational pattern contained in a WCC matrix have been suggested (Boker et al., 2002; Moulder, Boker, Ramseyer, & Tschacher, 2018), including the mean peak lag and the mean peak correlation. Other approaches for the association between time series include coherence (Bakhshayesh, Fitzgibbon, Janani, Grummett, & Pope, 2019; Codrons, Bernardi, Vandoni, & Bernardi, 2014), the Actor-Partner Interdependence Model (APIM; Fitzpatrick, Gareau, Lafontaine, & Gaudreau, 2016), and Granger causality (Dauwels, Vialatte, Musha, & Cichocki, 2010).

Information theory has provided a number of metrics by which the association between time series may be determined. The basic approach for these methods was developed by Shannon (1948) in his seminal paper on communication. Other researchers that followed, including Garner (1962) and Watanabe (1960), applied these ideas for use in psychology and other related fields. Information theoretic metrics have been used in the past to determine the predictability of time series, including music. The dynamic information measures proposed by Abdallah and Plumbley (2009) and Boker and Kubovy (1998) integrate information theory measurements of surprise and predictability with the moment-to-moment information provided by Markov chains to segment music signals into moments of surprise and moments of predictability.

Other approaches emphasize the segmentability of interpersonal behaviors. Methods

which segment interpersonal behaviors into moments of symmetry and non-symmetry are of particular interest because this suits the expected intermittent nature of symmetry. Abney, Paxton, Dale, and Kello (2014) distinguished behavioral matching from complexity matching, where segments of complex time-based behaviors can be separated from less complex segments using power law functions. They demonstrated that complexity matching is a relevant conversation coordination tool by finding that conversants in an affiliative conversation match the complexity of speech of their partner at larger timescales than those in a neutral conversation. The Pairwise Approximate Spatiotemporal Symmetry (PASS) algorithm (Sjobeck, Boker, Scheidt, & Tschacher, n.d.) is another approach to segmenting time series. Here, time series are segmented based on their shared pattern of association, such that any moment of time may or may not be indicative of symmetry.

Symmetry between Agents in a Network. Most studies of interpersonal symmetry have been restricted to dyads. The reason for this is often two-fold: because the dyad is a meaningful basic unit of larger groups, and because dyadic analyses are the simplest to implement. This restriction, while often practical and still insightful, is not strictly necessary. Similarity has long been theorized to be important at the group level (Codrons et al., 2014; Hanel, Maio, & Manstead, 2019; Simon, 1952). There is an apparent lack of research in group-level symmetry, which is arguably the result of a lack of methodology suited to capture association between three or more agents. Recently, however, there has a been a push to tackle this complicated problem. Gordon et al. (2020), for instance, found a relationship between synchrony and group cohesion in a trio of drummers. Participants were asked to drum along to either a predictable or unpredictable musical pattern. Participants exhibited more synchrony in heart rate during the task than at baseline, which suggests that the shared act of drumming influenced the physiological synchrony between each drummer. Those who drummed to a predictable beat reported higher group cohesion than those who drummed to an unpredictable beat. Finally, synchrony in heart rate across the triad predicted group cohesion. These results suggest

that symmetry is a relevant group-level phenomenon, and that further work is needed to explicate these relationships. Therefore, an algorithm which segments behavioral or physiological signals from interacting people into moments of symmetry and non-symmetry may be useful in understanding group-level phenomena like cohesion.

The human brain can likewise be thought of as a network. In this way one might ask about the functional connectivity of brain regions. Regions of the brain which are functionally connected transmit signals to each other in order to maintain communication. Here, as well, redundancy in the transmitted signal has communicative properties. Pillai and Jirsa (2017) suggest that the connectivity in a brain network is maintained by symmetry and that symmetry breaking is a functional entity in neural mechanisms. An algorithm which segments signals from related brain regions into moments of symmetry and non-symmetry may then be useful in understanding the connectivity of these brain regions.

The goal of the TASS algorithm is to segment three time series into moments of symmetry and non-symmetry. The sections that follow give further explanation to three particularly relevant aspects of the threeway symmetry segmentation problem: first, a mathematical discussion on the nature of symmetry; next, a discussion of information theory, the basis for the association metric used in the TASS algorithm; and last, a discussion of the segmentation problem and the basis on which segmentation will be tackled here.

Mathematical Symmetry

Symmetry is a ubiquitous phenomenon in nature, and is therefore a highly useful mathematical idea. In fields with complicated moving parts that are constantly changing, symmetry provides the ability to measure that which remains the same. In doing so, patterns are found, giving way to a structural understanding that can simplify problems. Symmetry analyses have been used to great effect in fields like chemistry (Reutzel-Edens, 2020; Wagner & Ashkenasy, 2009; Zabrodsky, Peleg, & Avnir, 1992), crystalography (Le

Page, 1987; Moeck, 2019; Spek, 2003), geology (Choukroune, Gapais, & Merle, 1987; Oglesby, Archuleta, & Nielsen, 1998; Saspiturry et al., 2019), and physics (Gross, 1996; Michel, 1980; Pais, 1966) to explicate the structures of similarity in physical systems.

Symmetry can be considered within one of three broad categories: *spatial symmetry*, *temporal symmetry*, or *spatiotemporal symmetry*. Spatial symmetry includes all symmetry that occurs across some distance and without consideration to time. Temporal symmetry involves similarity across time, without consideration to space or distance. Spatiotemporal symmetry includes any symmetry which is maintained in both space and time concurrently. Intersignal symmetry is given as a kind of spatiotemporal symmetry, since it includes both a spatial and a temporal component. The spatial component occurs in the equivalence of the behavior, as captured by the signal. The temporal component occurs in how the equivalence of behavior is maintained across time.

Under these broad domains, there are subcategories of symmetry forms. With each symmetry form, equivalence is maintained even after some transformation. Types of spatial symmetry include: *translation*, *rotation*, *reflection*, *glide*, *screw*, and *scale*. While any of these symmetry forms might be present in cases of intersignal symmetry, perhaps the most salient is reflection, or mirror, symmetry. Temporal symmetry forms include: *temporal displacement*, *temporal reflection*, and *temporal scale*. Similar to the spatial symmetry forms, the most salient temporal symmetry form is perhaps temporal reflection symmetry. Spatiotemporal symmetry forms takes from both spatial symmetry forms and temporal symmetry forms as needed.

Symmetry is often defined mathematically using several useful terms, which will be explored here. These terms are broadly defined so as to apply to any potential object or phenomenon of interest. First, symmetry is present in a *system*. That system is comprised of a number of *states*, or conditions of the system. The *state space* of a system refers to the collection of all possible states of the system. Any noteworthy subset of the state space may be referred to as a *subspace*. In order for a system to exhibit symmetry, any two states must

satisfy an *equivalence relation*, wherein two states are deemed equivalent by some criteria. The states which abide by this equivalence relation make up the *equivalence subspace*.

For example, assume we wish to mathematically describe the spatial symmetry of a perfect circle. Our system would be the circle, which is composed of a number of points. These points would be considered states, and the collection of all points would be the state space of our system. The criteria for two points of a circle to be equivalent and exist in the equivalence subspace is for those two points to be in the same spatial position. The goal, then, would be to find the noteworthy points on the circle which are contained within the equivalence subspace.

By defining our states as points on the circle, we are able to make claims about the rotational symmetry of circles. First, though, it is necessary to be able to capture the relationships between states. A *transformation* describes this relationship as the mapping of every state in the state space to some other state in the state space. If we wish to understand the rotational symmetry of the circle, then a rotation transformation is needed. Here, all points on the circle may be rotated by a constant angle about a fixed point (i.e., the center of the circle). In this way, we now have two circles: that which existed prior to the transformation, and that which existed after.

In order to define a *symmetry transformation*, it must be determined that the chosen transformation maps each state contained within the equivalence subspace to another state in the equivalence subspace. In so doing, the states of the equivalence subspace are shown to be invariant to the given transformation. In the circle example, the rotation transformation maps each point on the circle to an equivalent point on the circle, leaving the circle appearing unchanged. In other words, the circle that existed prior to the transformation is the same circle as the one created by the transformation. Thus we can say that a circle exhibits rotational symmetry because it demonstrates invariance to rotational transformations.

Let's consider an example of dyadic symmetry, where symmetry is described between

two objects. Here the system is two light switches which each have two positions, on and off. The state space would include the following: (1) the case where both light switches are turned on; (2) the case where the first light switch is on, and the second light switch is off; (3) the case where the second light switch is on, and the first light switch is off; and (4) the case where both light switches are turned off. The transformation maps the state of one light switch to the state of the other light switch. It is evident here that symmetry in the position of the light switches is present in case (1) and case (4). These two cases, then, represent a symmetric subgroup of the state space which is invariant to the transformation.

This same rationale may be applied to the spatiotemporal symmetry between more than one source. Consider two people engaged in a conversation, where they exhibit spatiotemporal symmetry in their head nod behavior. In the spatial domain, the heads of these two participants may exist in a number of spatial positions. A spatial transformation exists between the head position of one person and the head position of the other. The spatial transformation occurs for every instance in time, and so the temporal transformation considers the relationship between the spatial transformation of one time point and the spatial transformation of the successive time point. Note that the spatial transformation and the temporal transformation cannot be considered separately since both occur concurrently. Thus spatiotemporal symmetry is only present in the head nods of these conversants if both of the following conditions are met: (1) the spatial transformation maps the head position of one person to an equivalent head position of the other person; and (2) the temporal transformation maps the equivalence between head positions across successive time points. In this way, it can be said that the symmetry between head position is maintained across time.

Approximate Symmetry. So far the discussion of mathematical symmetry has remained in the realm of exactness. While exact symmetry may be present in your morning interactions with the mirror, it is not common in practice between two autonomous agents. Thankfully there is a system in place for discussing intermediate

amounts of symmetry and fluctuations in the amount of symmetry across time. This kind of symmetry is called *approximate symmetry*.

Approximate symmetry allows for a range of equivalence to be specified, rather than an exact matching of states. To provide this range of equivalence, a distance from exact equivalence must be specified which indicates the furthest from 0 distance two states can be and still be considered equivalent (Rosen, 1983). This distance acts as a threshold, where values below this distance are considered symmetric and values above this distance are considered nonsymmetric. This relationship is represented mathematically with the following equation (Rosen, 1983):

$$d(u, T(u)) \leq \epsilon \tag{1}$$

where u is the original state, $T(u)$ is a transformation of the original state, $d(u, T(u))$ is the distance between u and $T(u)$, and ϵ is the distance threshold for approximate symmetry. All examples of symmetry in human systems will be assumed to be approximate in nature.

Information Theory

Association. The goal of any empirical symmetry algorithm is to measure the equivalence between any given states. This can be done using an association metric. Here association metrics can be used to determine the extent to which two or more psychological agents align their behavior. Perhaps the most common association metric is the product-moment correlation,

$$r_{xy} = \frac{\sum_i (x_i - \overline{x})(y_i - \overline{y})}{N s_x s_y} \tag{2}$$

where r_{xy} is the correlation between variables X and Y, x_i is an element of X, y_i is an element of Y, $\overline{x}$ is the mean of all elements in X, $\overline{y}$ is the mean of all elements in Y, N is the number of elements in both X and Y, and s_x and s_y are the standard deviations of X and Y, respectively.

Product-moment correlation is the basis for several symmetry algorithms, including cross-correlation, windowed cross-correlation, and the PASS algorithm. However, there are two general drawbacks to using product-moment correlation: 1) product-moment correlation captures linear relationships, which may not be the relationship of interest; and 2) product-moment correlation is designed to measure the association between two sets of data, and cannot be used to capture the association between three or more sets of data. This latter reason necessitates a different association metric for the current situation. Association metrics from the field of Information Theory have been used to capture association between three or more sets of data (Anastassiou, 2007; Boly et al., 2012; Schneidman, Still, Berry, & Bialek, 2003). Here, association is conceptualized in a way that is different from product-moment correlation. Product-moment correlation measures the extent to which values from one variable correspond linearly to values in a second variable. Information theory association metrics, however, measure the extent to which one variable reduces the uncertainty in a second variable, in the bivariate case, or a set of variables in the multivariate case. Total correlation (Watanabe, 1960), an information-based association metric, will be used in the TASS algorithm to capture the association between three time series. The following section will elucidate the context within which total correlation was created and should be interpreted, starting with an explanation of the theoretical communication paradigm.

The Communication Paradigm. Much of psychological research is structured to answer questions about possible influence between measured variables. Clinicians, for example, may be interested in understanding whether an individual's anxiety level can affect their level of depression. It is often difficult, however, to establish influence between variables in empirical studies.

Communication theory was developed by Shannon (1948) to provide a structure in which the transfer of information could be understood. The structure is inherently broad and makes no assumption about the content of the information. Instead, the focus is on

probabilistic relationships between a set of possible outcomes. In this way, a message is considered informative if it occurs despite low probability.

It is perhaps obvious to suggest that the communication paradigm can be applied to interpersonal conversation, since interpersonal conversation is a primary way that information is transferred between people. However, since communication is broadly defined by the paradigm, it is not necessary to restrict its use to this situation. Instead, it is useful to think of communication as occurring between any two agents that have information transferred between them. In this way, brain regions or emotion states may be thought to communicate with each other.

The classic theoretical model for a discrete communication channel, established by Shannon (1948), gives us a structure within which we can conceptualize the relationships between communication sources. The theoretical model, shown in Figure 1, is meant to be general enough to capture the behavior of communication systems of any conceptual nature.

The flow of information for a discrete communication channel begins at a source and ends at a destination. While these two agents might both transfer information to each other, the channel is constructed for simplicity such that only one direction of information flow is considered at a time. This information must then travel across some contextually-appropriate distance, or channel. The distance is an integral part of the communication system, given that the two agents separated by zero distance are, in communication terms, indistinguishable and therefore cannot communicate. In communication systems it may be that this channel is subject to noise. Noise can alter information in transit such that the received message is not identical to the one that was sent. In order to be efficiently sent across the channel, the information must be processed into a parsimonious form and transmitted. A receiver, in connection with the destination, receives the information on the other side of the channel and processes the message in a form readable by the destination. Therefore, it is also possible that error exists in the

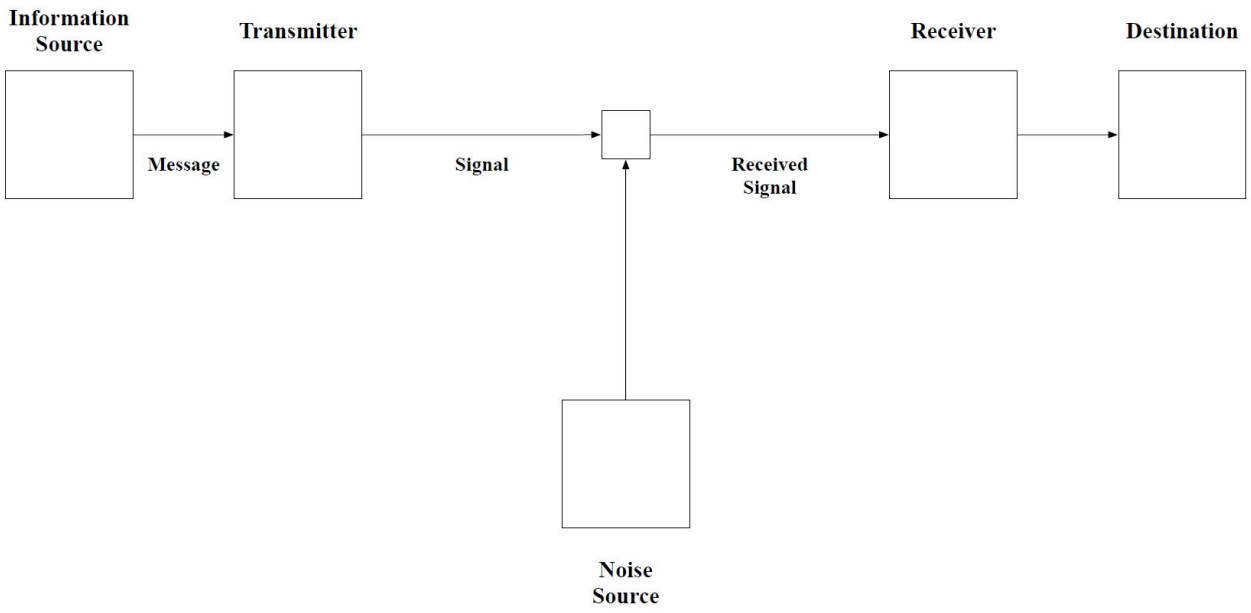

Figure 1. A discrete communication channel, as theorized by Shannon (1948). All circumstances where discrete information is transmitted from one destination to another can be conceptualized using this basic structure. A message, whose origin is the information source, is translated into a signal which is transmitted through a channel to a receiver. The receiver decodes the signal into a message, which can be interpreted at the destination. Noise exists in the channel, such that the message received by the destination may not be the one sent by the information source. This basic structure is the basis for information theory metrics of association.

transmitter or receiver or both, such that the information transmitted is not received in the intended way.

It is possible to conceptualize information transfer as occurring at different levels of experience. Consider two individuals engrossed in a game of charades, where the agent of acting is communicating, through gestures, an idea to the agent of guessing. The actor and the guesser in this example constitute the source and the destination, respectively. Thus, it may not be needed to specify the transmitter and the receiver in the game of charades, if one assumes that the actor is both the information source and the information transmitter

and that the guesser is both the information receiver and the information destination. However, it may be useful to delineate these elements of the communicating system with respect to other levels of human behavior. In this way, it may be more accurate to say that the cognitive system of the actor is the information source and the cognitive system of the guesser is the destination, where the actor's gesturing behavior is the transmitter and the visual perceptual system of the guesser is the receiver.

The influence of noise in this theoretical communication system is seen in the accuracy of the intended message as it is received by the destination. As per the natural circumstances in charades, it is possible for the actor to be participating in a behavior that they feel is clearly representative of the intended idea, but which is to some degree lost on the guesser. In this case, the miscommunication may be because of poor comprehensibility of the act, by way of the actor, or poor inference of the meaning of the act, by way of the guesser. These two possibilities represent noise in the transmission and translation of the information, respectively. Environmental noise may be a third source of disruption in the accurate communication of information from the source to the destination. In our game of charades, this could be anything from a distracting loud noise to the lighting in the room.

Another phenomenon not captured by the two-agent one-message simplicity of the discrete communication channel is the complexity of the transmitted signal. Thus the signal may contain multi-dimensional information that makes the interpretability of the signal more difficult at the destination. This would be akin to the actor producing a behavior which did not represent the simplest aspect of the idea, and thus did not directly suggest the intended idea. In each instance, noise increases the likelihood that the destination receives the transmitted information with less accuracy than was intended.

With so many possible sources of noise, it is of major interest to communication researchers to determine how likely the message that is received at the destination is the one transmitted by the source. Information theory metrics were designed in this context to measure how similar the transmitted message is to the message received. These metrics

have since been used more broadly to measure the reduction in uncertainty between co-occurring variables which are not directly related to communication (Bakhshayesh et al., 2019; Boly et al., 2012; Honey, Kötter, Breakspear, & Sporns, 2007).

Uncertainty in a Single Variable. Information theoretic metrics for association share a similar approach to that of product-moment correlation. With correlation, it is necessary to understand how each of two measured variables vary individually and also how much they vary together. This holds true for association metrics in information theory, except the variability is measured not by variance, but by entropy.

Entropy reflects the average uncertainty of a probability distribution. The entropy of a single flip of a coin, for instance, is *greatest* when the probability of getting a heads and the probability of getting a tails are equivalent, i.e., when $p(H) = p(T) = 0.5$. This is to say, when the probability of heads and tails are equivalent, we have the *most uncertainty* about what the outcome of the coin flip will be. Alternatively, a coin with two heads or two tails has no uncertainty in its outcome, and therefore in this case no entropy is present.

The probability distribution can be discrete or continuous. Discrete probability distributions are used to determine discrete entropy, and continuous probability distributions are used to determine the limiting density of discrete points (Jaynes, 1963, 1968), the continuous version of discrete entropy (not to be confused with differential entropy; Shannon, 1948). Discrete entropy will be used here. Discrete entropy can be calculated as such:

$$H(X) = -\sum_{x} p(x) \log p(x) \tag{3}$$

where $H(X)$ is the discrete entropy of random variable x and $p(x)$ is the probability of a given categorical outcome of x. The base of the logarithm determines the scaling, but not the interpretation, of the entropy value. Common logarithm bases include: 2, where the units are called bits; 10, where the units are called hartleys; and e, where the units are called nats. The current study will use base 2. All formulas presented here will omit the

base in acknowledgement that any could be used.

Empirical entropy can be calculated for empirically-collected data. This requires that the probability distribution be derived from the measured values in the data. It is common practice to discretize the probabilities associated with each measured value, which is to say that each value is assigned to a range, and that range is given a probability that is used in the entropy calculation. This process is called binning, and is similar to the binning process found in histograms. Empirical entropy will be used in this study.

Bivariate Information Transfer. Our goal is to capture relationships between sets of data. The bivariate case is the simplest of these cases. Bivariate relationships are present when the occurrence of an outcome in one variable provides information about the occurrence of an outcome in the other variable. Consider the Venn diagram in Figure 2. Here, the red circle corresponds to $H(X)$, the entropy of variable X, and the yellow circle corresponds to $H(Y)$, the entropy of variable Y. A relationship between X and Y is present when these two circles overlap. This relationship is often called *mutual information*. In this section, we will describe three possible calculations for mutual information: one which uses the joint entropy between X and Y, one which uses the conditional entropy between X and Y, and one which uses a direct formula.

The first formulation for mutual information is intuitive because it compares the maximum amount of uncertainty to the amount of uncertainty present between X and Y given their overlap. The maximum amount of uncertainty can be thought of as the sum total of uncertainty present in variables X and Y assuming no overlap, and is calculated by adding together $H(X)$ and $H(Y)$. The latter source of entropy, often called joint entropy, is the entropy of the joint event of outcomes in two variables. We know that each outcome in X is associated with a probability, $p(X = x_i)$, and that each outcome in Y is associated with a probability, $p(Y = y_i)$. If X co-occurs with Y, then each outcome in X has a probability of co-occurring with each outcome in Y. This is referred to as the joint probability, $p(X = x_i, Y = y_j)$. A sample joint probability matrix can be found in Figure 3.

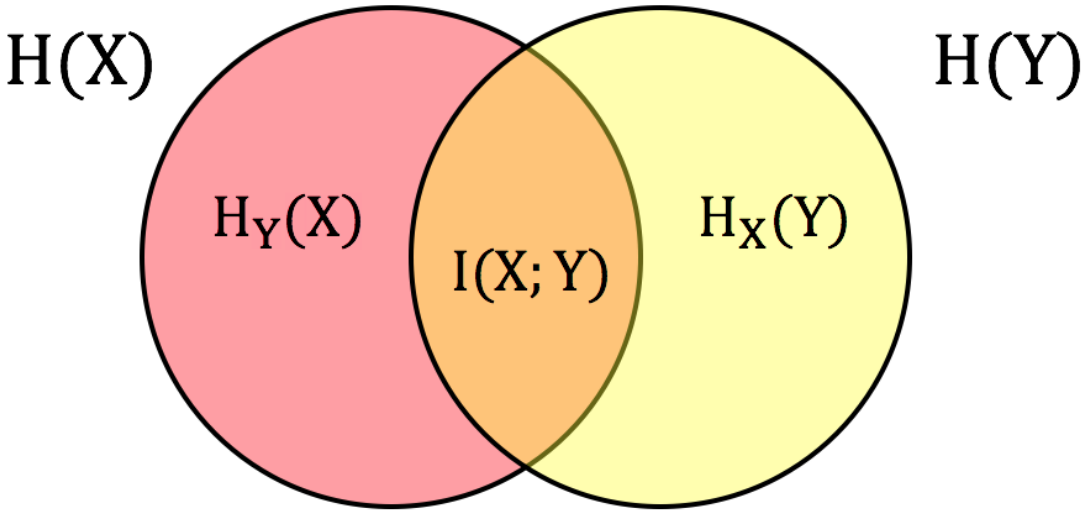

Figure 2. A Venn diagram representation of the entropy relationships between two variables. The red circle represents the entropy in variable X and the yellow circle represents the entropy in variable Y. Their overlap, the orange region, represents the reduction in uncertainty that each provides for the other. This region is the region of interest.

Joint entropy describes the uncertainty of the joint event (X, Y). The formula for joint entropy is as follows:

$$H(X,Y) = -\sum_{x}\sum_{y} p(x,y)\log p(x,y) \tag{4}$$

where $H(X,Y)$ is the joint entropy of variables X and Y, and $p(x,y)$ is the joint probability of X and Y. In the Venn diagram, joint entropy corresponds to the area of both circles given their potential overlap.

To calculate the amount of overlap, which corresponds to mutual information, the following formula may be used:

$$I(X;Y) = H(X) + H(Y) - H(X,Y) \tag{5}$$

X \ Y	Y = 1	Y = 2	Y = 3	Y = 4	Y = 5	
X = 1	p(X = 1, Y = 1)	p(X = 1, Y = 2)	p(X = 1, Y = 3)	p(X = 1, Y = 4)	p(X = 1, Y = 5)	p(X = 1)
X = 2	p(X = 2, Y = 1)	p(X = 2, Y = 2)	p(X = 2, Y = 3)	p(X = 2, Y = 4)	p(X = 2, Y = 5)	p(X = 2)
X = 3	p(X = 3, Y = 1)	p(X = 3, Y = 2)	p(X = 3, Y = 3)	p(X = 3, Y = 4)	p(X = 3, Y = 5)	p(X = 3)
X = 4	p(X = 4, Y = 1)	p(X = 4, Y = 2)	p(X = 4, Y = 3)	p(X = 4, Y = 4)	p(X = 4, Y = 5)	p(X = 4)
X = 5	p(X = 5, Y = 1)	p(X = 5, Y = 2)	p(X = 5, Y = 3)	p(X = 5, Y = 4)	p(X = 5, Y = 5)	p(X = 5)
X = 6	p(X = 6, Y = 1)	p(X = 6, Y = 2)	p(X = 6, Y = 3)	p(X = 6, Y = 4)	p(X = 6, Y = 5)	p(X = 6)
	p(Y = 1)	p(Y = 2)	p(Y = 3)	p(Y = 4)	p(Y = 5)	

Figure 3. The joint probability matrix between variables X and Y. Here, variable X has six outcomes, $\mathcal{X} = \{1, 2, 3, 4, 5, 6\}$, and variable Y has five outcomes, $\mathcal{Y} = \{1, 2, 3, 4, 5\}$. Each element of this matrix represents the joint probability of simultaneous outcomes in X and Y. The marginal probabilities here represent the summation of all probabilities in the associated row or column, and is equivalent to the probability of one outcome in either X or Y without consideration to the other variable or any of its outcomes. These joint probabilities are used to calculate the joint entropy between two variables.

where $I(X;Y)$ is the mutual information between variables X and Y, $H(X)$ is the entropy of X, $H(Y)$ is the entropy of Y, and $H(X,Y)$ is the joint entropy of X and Y. The sum $H(X) + H(Y)$ represents the maximum entropy that can exist in this system, and corresponds to the case where $H(X)$ and $H(Y)$ do not overlap, i.e., do not provide information about each other. If $H(X)$ and $H(Y)$ do not overlap, then the joint entropy $H(X,Y)$ will be the same size as the maximum entropy, and $I(X;Y) = H(X) + H(Y) - H(X,Y) = 0$. If $H(X)$ and $H(Y)$ completely overlap, and $H(X)$ and $H(Y)$ are the same size, then the maximum entropy will be exactly double the joint entropy.

The second formula for mutual information uses conditional entropy. Conditional entropy is the uncertainty in one variable given that the other variable has an outcome that is known. Conditional entropy is often calculated using the joint probability matrix, since the joint probability matrix contains conditional information in addition to joint information. The following is the formula for conditional entropy:

$$H(Y|X) = -\sum_{x,y} p(x,y) \log \frac{p(x,y)}{p(x)} \tag{6}$$

where $H(Y|X)$ is the conditional entropy of Y given X, $p(x,y)$ is the joint probability of X and Y, and $p(x)$ is the probability of X. This gives the area of the Venn diagram that is uniquely associated with $H(Y)$, removing the region where $H(X)$ overlaps with $H(Y)$. In order to calculate the conditional entropy of X given Y, or $H(X|Y)$, $p(y)$ is used in the denominator of the logarithm term instead of $p(x)$.

Mutual information can be calculated from conditional entropy using the following calculation:

$$\begin{aligned} I(X;Y) &= H(X) - H(X|Y) \\ &= H(Y) - H(Y|X) \end{aligned} \tag{7}$$

where $I(X;Y)$ is the mutual information between X and Y, $H(X)$ is the entropy of X, $H(Y)$ is the entropy of Y, $H(Y|X)$ is the conditional entropy of Y given X, and $H(X|Y)$ is the conditional entropy of X given Y. Here, the overlapping region is determined by taking the entropy of one variable and subtracting the entropy unique to that variable to produce the entropy which is shared by both variables. This can be done using either variable.

Lastly, the mutual information can be calculated directly using the following formula:

$$I(X;Y) = \sum_{x,y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)} \tag{8}$$

where $I(X;Y)$ is the mutual information between X and Y, $p(x,y)$ is the joint probability between X and Y, $p(x)$ is the probability of X, and $p(y)$ is the probability of Y. Here, the denominator of the logarithm contains both $p(x)$ and $p(y)$. It should be emphasized that all three formulations produce the same value for mutual information. While mutual information is the most common bivariate association metric in information theory, it may be useful to consider the joint entropy or the conditional entropy in some specific cases.

Trivariate Association. A metric which measures the association between three or more variables is required for capturing moments of threeway symmetry. A number of useful sources of entropy between three or more variables exist, any of which represent a different structural approach to measuring association. The most common of these is total correlation (Watanabe, 1960) (also called total constraint; Garner, 1962). The TASS algorithm will use total correlation as its trivariate association metric, but this section will cover several possible candidates.

Consider Figure 4, which depicts the entropy relationships between three variables X, Y, and Z. The desired quantity, total correlation, is the sum-total of all overlap between the three circles, $C(X;Y;Z)$. This quantity is not directly pictured in the diagram; however, it is present in the following relationship:

$$C(X;Y;Z) = H_Y(X;Z) + H_Z(X:Y) + H_X(Y:Z) + 2H(X;Y;Z). \quad (9)$$

These quantities and their derivations will be explained below. Note however that this metric does not require the three circles to overlap, as association will be found when any two circles overlap. In the case where no trivariate overlap is present, the total correlation is the sum of the pairwise relationships. This entity is the most commonly used because of its straightforward interpretation and the fact that it naturally extends from mutual information.

Total correlation is a multivariate extension of mutual information which can be used to measure the association between three variables, and can therefore be calculated in a similar way. To simplify calculations, only the joint entropy approach to calculating total correlation will be given. First, it is necessary to measure the maximum entropy between all three variables, calculated as the sum of the univariate entropies of each variable:

$$H_{max}(X,Y,Z) = H(X) + H(Y) + H(Z) \quad (10)$$

where $H_{max}(X,Y,Z)$ is the maximum entropy between variables X, Y, and Z, $H(X)$ is the entropy of X, $H(Y)$ is the entropy of Y, and $H(Z)$ is the entropy of Z. The maximum entropy constitutes the greatest possible amount of entropy between all three variables, which occurs when no variable overlaps with any other variable.

To calculate the joint entropy of three variables, it is useful to begin with the *multiple conditional entropy*, the multivariate extension of conditional entropy in two variables. Multiple conditional entropy measures the unique entropy in a variable, given the occurrence of the other two variables. In the Venn diagram in Figure 4, $H_{XY}(Z)$, $H_{XZ}(Y)$, and $H_{YZ}(X)$ are all multiple conditional entropy sources. The formula for the multiple conditional entropy of Z given X and Y, $H_{XY}(Z)$, is given as:

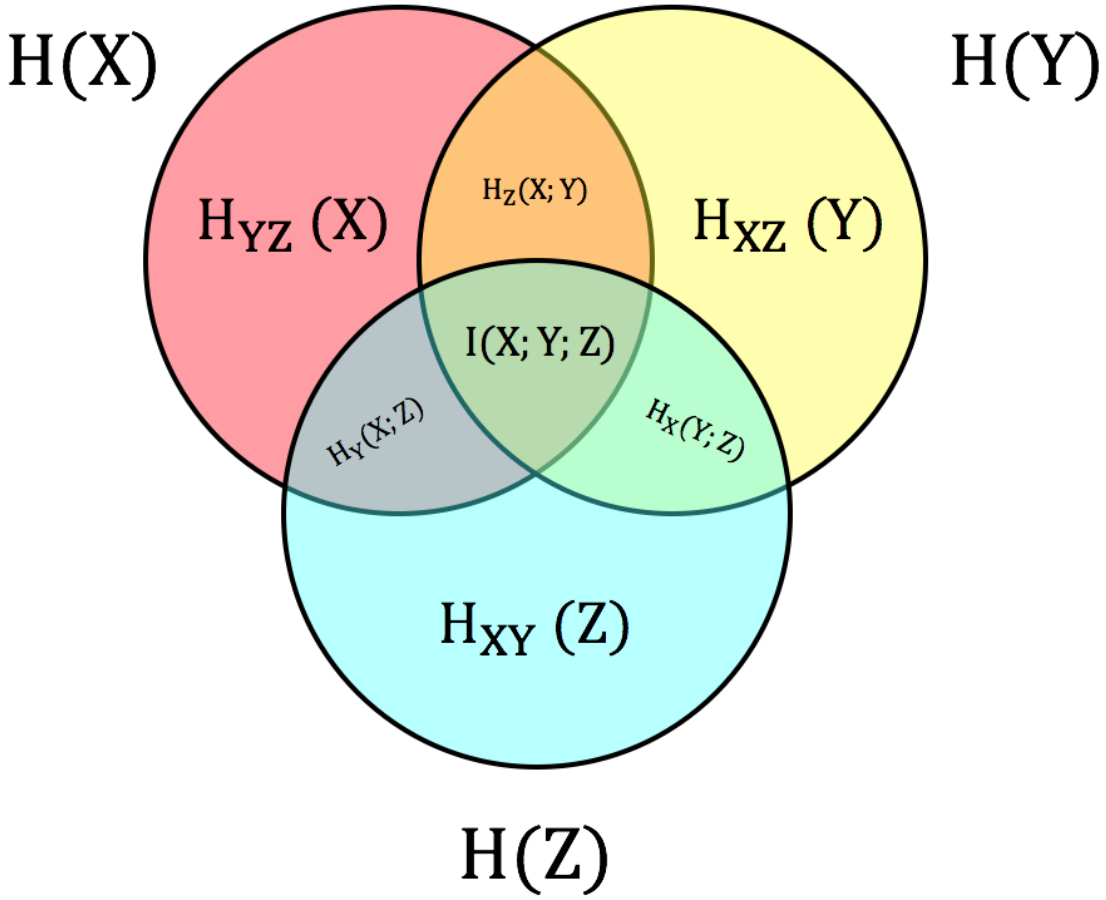

Figure 4. A Venn diagram representation of the entropy relationships between three variables. The red circle represents the entropy in variable X, the yellow circle represents the entropy in variable Y, and the blue circle represents the entropy in variable Z. Their overlap represents the reduction in uncertainty that each provides for the other. This region is the region of interest.

$$H_{XY}(Z) = -\sum_{x,y} p(x,y) \sum_{y} p_{xy}(z) \log p_{xy}(z) \tag{11}$$

where $H_{XY}(Z)$ is the multiple conditional entropy of Z given X and Y, $p(x,y)$ is the joint probability of X and Y, and $p_{xy}(z)$ is the conditional probability of Z given X and Y. This will be used to calculate the joint entropy between three variables.

The joint entropy captures the uncertainty jointly present in all three variables together, given the fact that they reduce uncertainty in each other. Joint entropy, $H(X,Y,Z)$ is calculated as follows:

$$H(X,Y,Z) = H(X) + H_X(Y) + H_{XY}(Z) \tag{12}$$

where $H(X,Y,Z)$ is the joint entropy between variables X, Y, and Z, $H(X)$ is the univariate entropy of X, $H_X(Y)$ is the conditional entropy of Y given X, and $H_{XY}(Z)$ is the multiple conditional entropy of Z given the occurrence of both X and Y. Notice that these three regions cover the entirety of the colored area in the Venn diagram in Figure 4, and only count the overlapping regions once.

Total correlation, $H(X;Y;Z)$, can be calculated using maximum entropy and joint entropy in the following way:

$$H(W;X;Y) = H_{max}(X,Y,Z) - H(X,Y,Z) \tag{13}$$

where $H(X;Y;Z)$ is the total correlation between variables X, Y, and Z, $H_{max}(X,Y,Z)$ is the maximum entropy, and $H(X,Y,Z)$ is the joint entropy. This value represents the association between three variables and will be used accordingly in the TASS algorithm. However, there are other sources of entropy which represent other ways of capturing the association between three variables, and these will be discussed here.

Multiple contingent entropy (Garner, 1962) is the entropy that one variable shares with the other two variables. Multiple contingent entropy can be measured by subtracting the unique uncertainty in a variable from the overall uncertainty in that variable. What remains is the overlap between that variable and the circles representing the other two variables. This can be seen with the following formula:

$$H(Z; X, Y) = H(Z) - H_{XY}(Z) \tag{14}$$

where $H(Z; X, Y)$ is the multiple contingent entropy of Z given X and Y, $H(Z)$ is the univariate entropy of Z, and $H_{XY}(Z)$ is the multiple conditional entropy in Z given that X and Y have occurred. Note that the multiple contingent entropy and the multiple conditional entropy sum to the total entropy contained in Y. Even though this association metric takes into account the influence of all three variables, it is perhaps best used to answer questions about the relationship a particular variable has with other variables.

Partial contingent entropy (Garner, 1962) captures the relationship between two variables given the occurrence of the other variable. This can be computed from multiple contingent entropy in the following way:

$$H_X(Y : Z) = H(Z; X, Y) - H(X; Z) \tag{15}$$

where $H_X(Y : Z)$ is the partial contingent entropy between Y and Z given that X has occurred, $H(Z : X, Y)$ is the multiple contingent entropy of Z given that X and Y have occurred, and $H(X; Z)$ is the mutual information between X and Z. This relationship is similar in concept to partial correlation (Garner, 1962), and may be useful in cases where the relationship between two variables are of interest, and it is desired to partial out the influence of a third variable.

Dual total correlation (DTC; Garner (1962)) constitutes a different approach to the multivariate extension of mutual information. DTC is calculated in the following way:

$$D(X;Y;Z) = H(X,Y,Z) - H_{XY}(Z) - H_{YZ}(X) - H_{XZ}(Y) \tag{16}$$

where $D(X;Y;Z)$ is the DTC for variables X, Y, and Z, $H(X,Y,Z)$ is the joint entropy between all three variables, $H_{XY}(Z)$ is the multiple conditional entropy of Z given X and Y, $H_{YZ}(X)$ is the multiple conditional entropy of X given Y and Z, and $H_{XZ}(Y)$ is the multiple conditional entropy of Y given X and Z. Here is perhaps a more intuitive association metric which accounts for pairwise relationships but only accounts for trivariate relationships once. However, this approach will not be used.

The last alternative association metric that will be covered here is interaction information, or $I(X;Y;Z)$ (Garner, 1962). This is the pure trivariate relationship. However, unlike total correlation and DTC, interaction information can be negative. Consider the equation for interaction information,

$$\begin{aligned} I(X;Y;Z) &= I(X;Y) - H_Z(X;Y) \\ &= I(X;Z) - H_Y(X;Z) \\ &= I(Y;Z) - H_X(Y;Z) \end{aligned} \tag{17}$$

where $I(X;Y;Z)$ is the interaction information between variables X, Y, and Z, $I(X;Y)$ is the mutual information between X and Y, $I(X;Z)$ is the mutual information between X and Z, $I(Y;Z)$ is the mutual information between Y and Z, $H_Z(X;Y)$ is the partial contingent entropy between X and Y given Z, $H_Y(X;Z)$ is the partial contingent entropy between X and Z given Y, and $H_X(Y;Z)$ is the partial contingent entropy between Y and Z given X. If, in the first case, X and Y jointly account for uncertainty in Z but are independent when not accounting for Z, then $H_Z(X;Y)$ will be larger than $I(X;Y)$ and I(X;Y;Z) will be negative. This changes the interpretation of the value, and for this reason interaction information is not often used, and will not be used in the current study. Now that it is clear which association metric will be used, it is necessary to establish the

approach that will be taken to segment time series based on the moments of association between three time series.

Segmentation of Time

The behavior of a system can always be understood as occurring across time. Although time is assumed to be a continuously-occurring process, empirical studies which measure behavior across time must do so in discrete steps according to the sampling rate. In this way, a single behavioral measurement is associated with a discrete moment in time at which the behavior occurs. These measurements occur in succession, in the order in which they occur in time. Empirical questions may be asked of the set of measured behaviors, aggregated across time; these questions could be called "questions of magnitude". It is also possible, however, to ask about *when* a behavior occurs, and how the occurrence of one behavior in time relates to the behavior of its nearest neighbors; these are "questions of temporal occurrence". Here, the time information is maintained, and the relationships between behaviors is aggregated with reference to time. The current study will focus on the latter set of questions.

Non-Stationarity. The emphasis on temporal occurrence naturally assumes that a span of measurement may contain more than a single kind of behavior, a phenomenon often referred to as "non-stationarity". Many areas of research that regularly deal with time series data study systems which may not meet this criteria. These areas include psychology (Boker et al., 2011; Molenaar, 2004), economics (Mikosch & Starica, 2004; Nelson & Plosser, 1982), and biology (Ivanov et al., 1996; Turchin & Taylor, 1992), among many others. The behavior of systems in these fields are dynamic by nature. This dynamic nature is a result of the relationship that exists between behavior and changes in contextual circumstances, both internal and external. Internal circumstances related to distributional changes in behavior over time may include boredom (Barbalet, 1999) or anticipation (Magyari, Bastiaansen, de Ruiter, & Levinson, 2014). External circumstances may include changes in the behavior of others (Cappella, 1997; Schmidt & Richardson,

2008) or new information (Oullier, de Guzman, Jantzen, Lagarde, & Scott Kelso, 2008). It is not realistic to expect that the statistical characteristics of a given interval sampled from the phenomenon do not depend on when it was sampled.

Many statistical approaches to the measurement of symmetry, particularly cross-correlation, assume that the similarity behavior abides by a stationary process, an unfortunate circumstance given the large number of non-stationary similarity behaviors. Several ways of avoiding the assumption of stationarity have been suggested in the literature and will be discussed here. Hendry and Juselius (2000) suggest that the dynamic behavior of economic data could be better captured with stochastic trends rather than linear trends. They offer unit-rate processes as a possible solution—where linear sections of stochastic trends are considered individually and allow for the stochastic nature of the data to be maintained in its parameter set. The process of focusing on stochastic trends is sometimes present in symmetry methods, but this approach is not widely used.

The solution suggested by Boker et al. (2002) is to assume the local dynamics of the system behave in stationary ways. Whereas the assumption of global stationarity posits that a measured variable behaves with regularity for any given interval length, the assumption of local stationarity allows for a measured variable to behave consistently for a short specified interval of time but that the variable may exhibit different behaviors depending on the starting time. This modified assumption is present in their method for assessing the co-variation in lagged segments of paired time series, called windowed cross-correlation (WCC). The WCC method is used in the current method to capture behavioral coordination.

Global stationarity requires that the global parameters of the entire time series represent the behavior of the overall system and concurrently the behavior of any section of the time series of a given interval length. It is apparent, therefore, that time series exhibiting local stationarity must be comprised of segments of various lengths, where each segment corresponds to a set of invariant parameters. The dyadic case presents coupled

time series which may separately adhere to the local stationarity assumption. Each time series would then be described by a given number of segments. These segments could be determined individually, ignoring information from the other time series. Currently there exist methods for determining the proper lengths of individual segments which together comprise a single time series. Keogh, Chu, Hart, and Pazzani (2004) offer several methods which utilize data mining to determine which intervals of time series correspond to meaningful segments. Many of these methods make an assumption of linearity.

However, the claims made by these methods about the segmentation of time series are relevant to each individual time series and not to the dyadic relationship that exists between them. In order for the dyadic relationship to be segmented, a segmentation algorithm must operate on data that represents the dyadic relationship and not simply each of the relevant time series. Ideally, this dyadic relationship would be captured with a method that appealed to the potential nonstationary nature of the original time series. The Pairwise Approximate Spatiotemporal Symmetry (PASS) algorithm (Sjobeck et al., n.d.) is such an approach.

Pairwise Symmetry Segmentation. The Pairwise Approximate Spatiotemporal Symmetry (PASS) algorithm (Sjobeck et al., n.d.) was designed to segment pairwise time series based on the moments of symmetry and nonsymmetry that exist between them. The algorithm does not assume a particular nature for the time series, and for this reason the algorithm is applicable in a number of dyadic time series situations. Currently, the algorithm has been used to capture symmetry in the body movement of therapists and clients in therapy sessions (Sjobeck et al., n.d.), and the head movements of conversation partners.

The PASS algorithm relies on several notions which will briefly be discussed here. First, similarity between time series is captured with product-moment correlations, since this provides an estimate of the paired association between two sets of data. Correlation has a number of useful properties which correspond to our understanding of time series

association: a perfect positive correlation corresponds to perfect similarity, a perfect negative correlation corresponds to perfect dissimilarity, and zero correlation corresponds to no association. To meet a local stationarity assumption, these correlations are taken between windows of the time series, rather than the full time series. Thus, multiple correlation values taken together represent the correlation pattern, rather than just a single value.

Should association be expected only between windows which occur over the same set of time points? It is assumed here that most psychological agents cannot predict the behavior of others well enough to expect simultaneous relationships. For this reason, windows are lagged to account for possible association that does not occur instantaneously. The resulting correlation pattern is contained in a matrix, with window location as one dimension and lag as the other. The Windowed Cross-Correlation (WCC) procedure (Boker et al., 2002) provides this correlation matrix in exactly the way specified. WCC allows the user to specify the following hyperparameters: w_{size}, the size of the window; w_{inc}, the number of time points forward a window is shifted; τ_{max}, the maximum lag value; and τ_{inc}, the number of time points backwards a window is lagged. The general process by which windows are lagged and shifted is evident in Figure 5.

Next, the PASS algorithm determines if the association pattern at every time point is similar to the association pattern of time points which succeed and precede each given time point. Successively similar association patterns are assessed using prospective correlations, and those which exist prior to the time point are assessed using retrospective correlations. Time points which share an association pattern with nearest time point neighbors are considered symmetric, as indicated by a distance threshold.

The goal of a symmetry segmentation algorithm is to determine when in time the behavior of two time series exhibit symmetry and when in time they do not. Thus, the algorithm assigns to each time point a dichotomous marker that indicates either "symmetry" or "nonsymmetry". When successive time points are collectively determined to

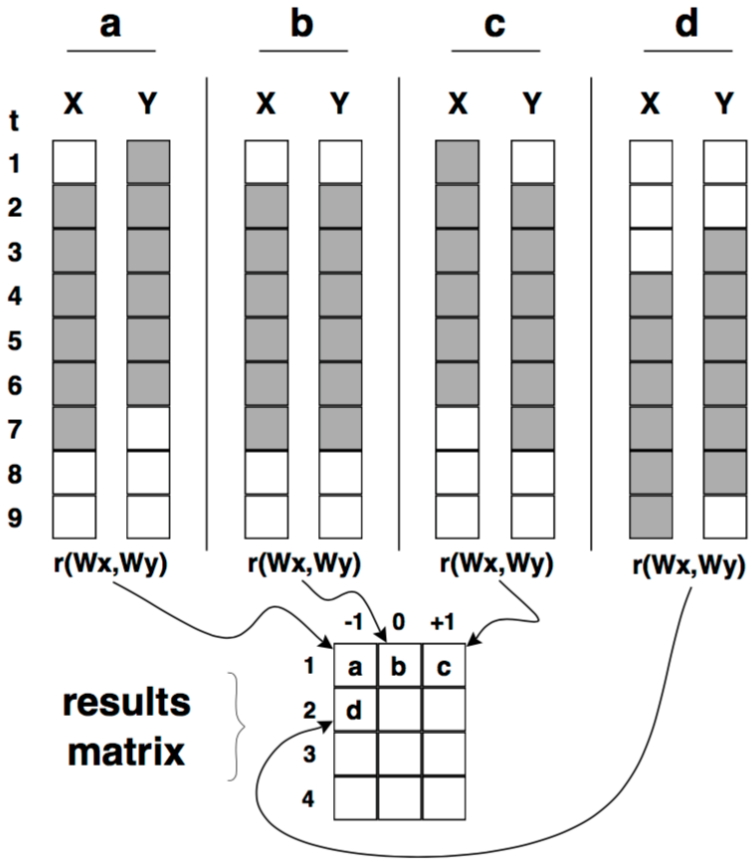

Figure 5. A diagram representation of the window process for windowed cross-correlation (WCC). The window size is 6 samples, the window increment is 2 samples, the maximum lag is 1 sample, and the lag increment is 1 sample. Note that (b) represents the default position, where neither window is lagged. The correlation for (b) is retained by the results matrix, at a lag position of 0 and at the first reference time point. (a) represents the situation where the window in Y is lagged, and (c) represents the situation where the window in X is lagged. (d) represents the case where the window reference point has been shifted down 2 in accordance with the window increment. Figure adapted from Boker et al. (2002).

be symmetric, a segment emerges. The first time point in this segment may be considered the "moment of symmetry formation". Likewise, the last time point in the segment may be considered the "moment of symmetry breaking". It is likely that the period of formation and breaking exists for a longer span of time than a single moment. Approaches which detect a formation and breaking period would then have four markers: symmetry formation, symmetry, symmetry breaking, and nonsymmetry. The PASS algorithm uses both a two-phase and a four-phase segmentation approach. The two-phase symmetry distinction is used to calculate symmetry metrics, while the four-phase symmetry distinction is used in the segmentation plots, like those in Figure 6.

The plots in Figure 6 are simulated examples which come from Sjobeck et al. (n.d.). Each example was chosen to provide insight into the utility of the PASS algorithm. Figure 6a suggests that the PASS algorithm can be used to determine if two sine curves have symmetric moments; in fact, the PASS algorithm determined that the entirety of these two time series were symmetric, which meets expectations. No symmetry was indicated in Figure 6b, the simulated noise example, which also meets our expectations. These two cases represent the two extreme cases for pairwise symmetry: perfect symmetry and no symmetry. Figure 6c demonstrates that the PASS algorithm can detect moments of symmetry among moments of non-symmetry with this segmented example. Notice that the sine curve segments, which we would expect to be symmetric, are indicated as symmetry, and the noise segments, which we would not expect to be symmetric, are indicated as non-symmetric. Figure 6d, however, demonstrates the limitations of the PASS algorithm by suggesting that poor symmetry indication is determine when segments are small and of varying length.

Threeway Symmetry Segmentation. Pairwise relationships are the most often considered when assessing the symmetry between two or more psychological agents. Group dynamics, like those represented in a network, may contain informative pairwise symmetry relationships, but it is likely that group-level symmetry phenomena are not simply the sum

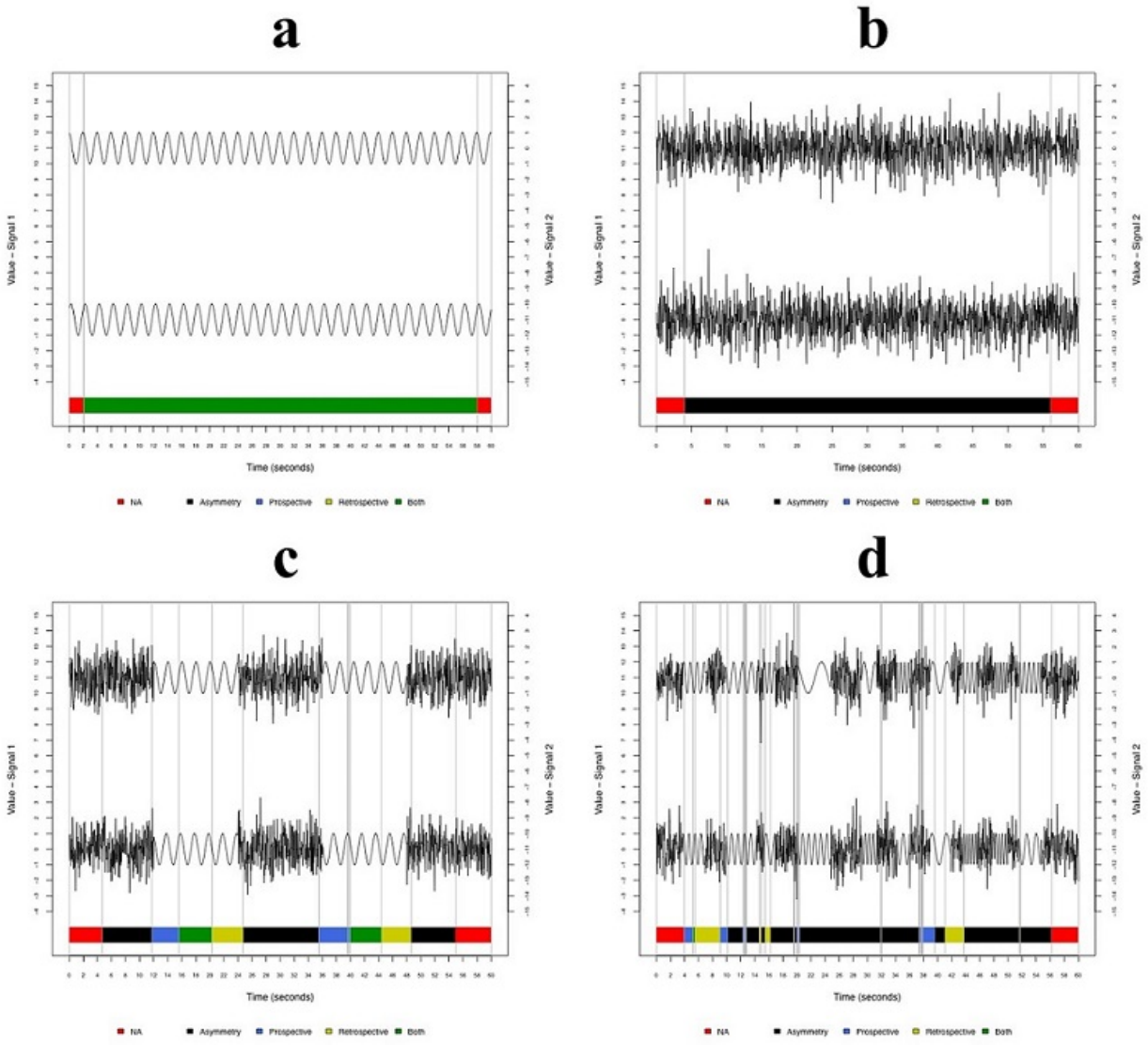

Figure 6. Segmentation plots of simulated examples which come from Sjobeck et al. (n.d.). Red segments indicate time points not included in the analysis; black represents time points indicated as "non-symmetric"; blue represents time points which exhibit only prospective symmetry relationships; yellow represents time points which exhibit only retrospective symmetry relationships; and green represents time points which exhibit both prospective and retrospective symmetry relationships. The following simulated data examples were chosen: (a) sine curves, which were considered perfectly symmetric; (b) simulated noise, which were considered non-symmetric; (c) segments of noise and sine curves, which were segmented into moments of non-symmetry and symmetry, respectively; and (d) segments of noise and sine curves of random length.

of their respective pairwise relationships. In other words, when three or more coordinating signals exhibit symmetry, pairwise relationships may be informative but lack insight into concurrent relationships between all signals. Consider the basic three-node network in Figure 7. Here, pairwise symmetry relationships are present as S_{XY}, S_{YZ}, and S_{XZ}. However, the symmetry relationship between all three nodes, S_{XYZ}, may provide insight into group symmetry dynamics that are not contained in the pairwise relationships. The goal of the TASS algorithm is to measure the symmetry between three signals, which can then be compared to the pairwise relationships to provide a complete picture of the symmetry dynamics between three signals.

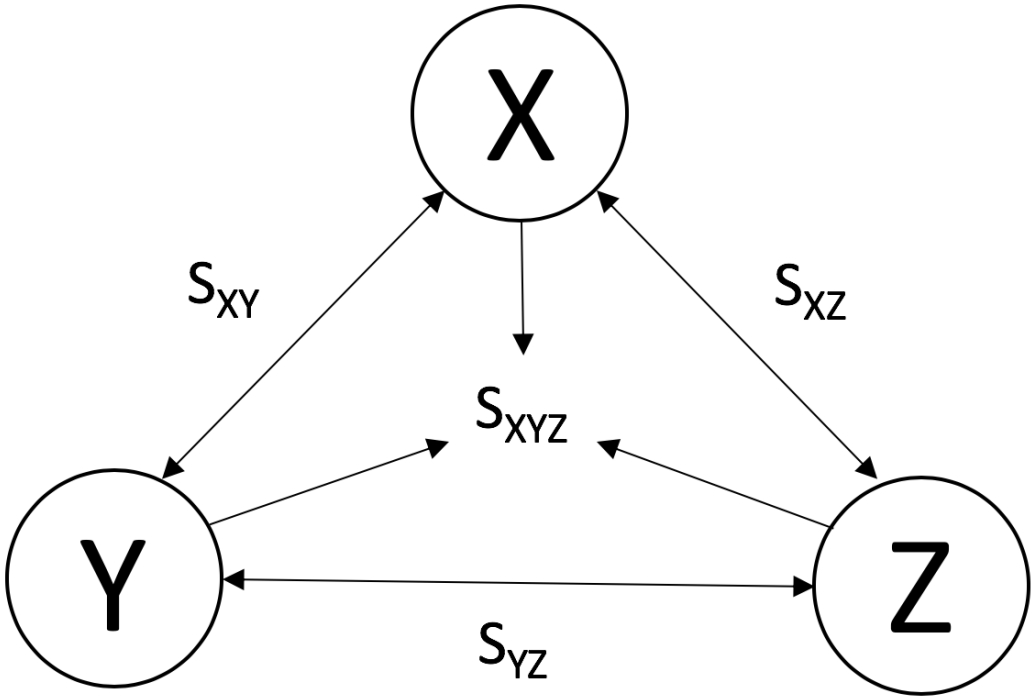

Figure 7. A basic three-node network between psychological agents X, Y, and Z. The symmetry values S_{XY}, S_{YZ}, and S_{XZ} represent the pairwise symmetry relationships between each pair of psychological agents. The symmetry value S_{XYZ} represents the threeway symmetry relationship between all three psychological agents.

A method which captures the symmetry in three or more signals is subject to many of the same considerations that a pairwise method might take into account: stationarity assumptions, lagged relationships, and segments of symmetry and non-symmetry, to name a few. For this reason, the TASS algorithm will extend the PASS algorithm to capture the symmetry between three signals. The goal of the current study is to assess the validity of the TASS algorithm with edifying simulations, and to determine the ecological validity of the algorithm using a number of empirical datasets.

The Threeway Approximate Spatiotemporal Symmetry (TASS) Algorithm

The current proposed method is an extension of the PASS algorithm to the analysis of symmetry in three time series, which will be referred to as the Threeway Approximate Spatiotemporal Symmetry (TASS) algorithm. Before discussing the algorithm and its steps, it is necessary to introduce the trivariate windowed cross-correlation (WCC-3), an extension of bivariate windowed cross-correlation to three signals. The WCC-3 uses total correlation instead of product-moment correlation to capture association between windows of the time series. For this reason, we will start with an exploration of total correlation.

Total Correlation

In the absence of a linear correlation metric which captures the association between three or more sources, we draw an association metric from the field of information theory. Several metrics exist and may be useful in capturing different kinds of symmetric relationships between three sources. We will use *total correlation* Watanabe (1960), which has in other literature been called *total restraint* or multivariate restraint Garner (1962). Total correlation with three sources is calculated in the following way:

$$C(X;Y;Z) = H_{max}(X,Y,Z) - H(X,Y,Z) \tag{18}$$

where $C(X;Y;Z)$ is the total correlation of sources X, Y, and Z, $H_{max}(X,Y,Z)$ is the

maximum entropy between the sources, and $H(X, Y, Z)$ is the joint entropy between the sources. Maximum entropy is the total amount of uncertainty in three measured sources, which does not take into account the interrelationships between them. Maximum entropy can be calculated as such:

$$H_{max}(X, Y, Z) = H(X) + H(Y) + H(Z) \tag{19}$$

where $H(X)$ is the entropy in source X, $H(Y)$ is the entropy in source Y, and $H(Z)$ is the entropy in source Z. The maximum entropy represents the system if all sources were orthogonal and therefore expressed no association. Joint entropy, on the other hand, is the *presented* uncertainty in all three sources. Joint entropy can be calculated in several ways, the following of which is just one way:

$$H(X, Y, Z) = H(X) + H_X(Y) + H_{XY}(Z) \tag{20}$$

where $H(X)$ is the uncertainty in X, $H_X(Y)$ is the uncertainty in Y given that X has occurred, and $H_{XY}(Z)$ is the uncertainty in Z given that X and Y have occurred. Notice that the terms which sum to the joint uncertainty are mutually exclusive sources of uncertainty, and they cover the entire area depicted in Figure 4.

Trivariate Windowed Cross-Correlation (WCC-3)

The first step in the TASS algorithm, trivariate windowed cross-correlation (WCC-3), captures the association between three time series in short lagged intervals. Short intervals are used in order to avoid avoid an assumption of global stationarity, i.e. stationarity across the entire time series. These intervals are lagged so that associations which do not occur simultaneously can still be captured by the TASS algorithm, as simultaneous association is assumed to be rare between autonomous agents. The time point at the beginning of each window in default position is referred to as the reference time point. This is the point in

time at which the TASS algorithm will make claims about the relationships between trivariate time series. The WCC-3 procedure is conceptually similar to the bivariate version of the windowed cross-correlation (WCC-2; Boker et al., 2002). Here, instead of product moment correlation, the total correlation between intervals of the three time series is taken.

Consider the following as a conceptual explanation for the WCC-3 procedure. An interval of chosen size is selected from each time series. These chosen intervals are referred to as windows. Each window begins in a default position which is the same across all three time series. The total correlation between these three windows in default position captures simultaneous relationships between the time series in the specified interval of time. Next, the windows are systematically lagged. When all lags are accounted for, the default position of the windows is incremented by a given amount, and total correlation is taken at different lags. These total correlation values are contained within a tensor of appropriate size. This process continues until the end of the time series is reached.

It is necessary to specify four hyperparameters: the size of the window, W_{size}; the maximum amount of lag, τ_{max}; the number of increments between lags, τ_{inc}; and the number of increments between default positions in the window, w_{inc}. These mirror the hyperparameters of the WCC-2 and can be interpreted in an equivalent way. W_{size} determines how many time points are contained in each window and thus how many data values are used to compute the total correlation. If the window is too small, then the accuracy of the total correlation value may be reduced. If the window is too large, then the total correlation value may aggregate too much of the data to be interpreted as a single process. Choice of a window size should depend on how quickly one expects the dynamics of a given time series to change. For instance, it is customary to use a window size of two seconds when analyzing conversation data (Boker et al., 2002).

Each window is lagged further and further backwards in time from the default position until a maximum lag is reached. τmax sets the maximum lag distance from the default position. A value for τ_{max} should be chosen as the largest amount of time a lagged

relationship might be expected. Too small a value for τ_{max} and the lagged relationship of interest may not be captured; too large a value for τ_{max} and spurious relationships between lagged windows may occur. The hyperparameter τ_{inc} determines how much a lagged window is shifted at each lag iteration. It is standard to set τ_{inc} to 1 sample.

w_{inc} determines how much the default position of the windows is shifted after the associations at all lags for the previous default position have been taken. It is standard for w_{inc} to be set to 1 sample, so that all time points in the data can be treated as a reference time point. This is particularly important for the TASS algorithm, since the TASS algorithm will make claims about each reference time point.

Figure 8 illustrates the WCC-3 process for a matrix slice of the resulting WCC-3 tensor. This matrix slice contains only lag information at a given reference time point. In this hypothetical example, the matrix is associated with the reference time point 6 samples. This is evident because the default position for the windows begins at sample 6. Here, each window has 5 samples (i.e., $w_{size} = 5$), the maximum lag is 5 (i.e., $\tau_{max} = 5$), and the lag increment is 1 (i.e., $\tau_{inc} = 1$). Each value of the matrix contains a total correlation value which relates windows in each of the three time series X, Y, and Z. The columns of this matrix are each associated with a different amount of lag in X. The rows of this matrix contain information about a lag position in both Y and Z. At the row marked 0, neither Y or Z are lagged from the default position. Rows above 0 lag the window in Z but not Y, and rows below 0 lag the window in Y but not Z.

Six examples of window lag positions are given. Figure 8e represents the case where all windows are at the default position. Figure 8d and Figure 8f represent the case where time series X is at the default position and either Z or Y, respectively, are at maximum lag. Note that when windows in either Z or Y are lagged, the window in the other time series is at default position. Figure 8b represents the case where Y and Z are at default window position, and X is at maximum lag. Figure 8a and figure 8c are similar to figure 8d and figure 8f, except now X is at maximum lag.

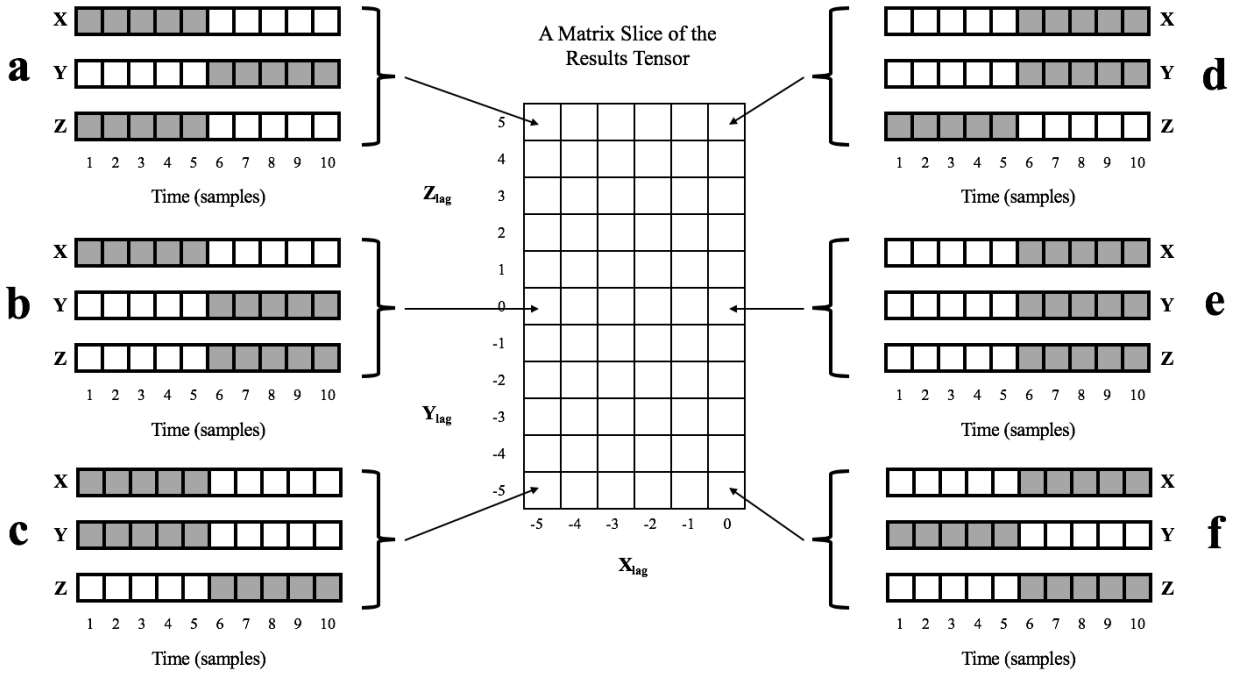

Figure 8. The trivariate windowed cross-correlation (WCC-3) process for producing a matrix slice of the results tensor associated with time point 6 seconds in this example. Here, the location of windows are lagged, and the association between these lagged windows are recorded in the matrix slice. Window orientation (e) represents the case where all windows are at a default position with no lag. Window orientations (a) and (d) represent the case where the window in time series Z is at the maximum lag away from the default position. Window orientations (c) and (f) represent the case where the window in time series Y is at the maximum lag away from the default position. Window orientations (a), (b), and (c) represent the case where the window in time series X is at the maximum lag away from the default position.

An example WCC-3 tensor is given in figure 9. It is evident from this figure that a window increment of 1 (i.e., $w_{inc} = 1$) was used because each time point in the time dimension increases by 1. Notice that the matrix created in Figure 8 is present at a time of 6 samples. Here we can see that each element of the matrix from figure 8 is followed in time by 14 total correlation values, where all elements have windows at the same lags relative to the reference time point. In this way the total correlation pattern at one time point is expected to evolve as the reference time point elapses. This information is crucial in the process of determining symmetry across time.

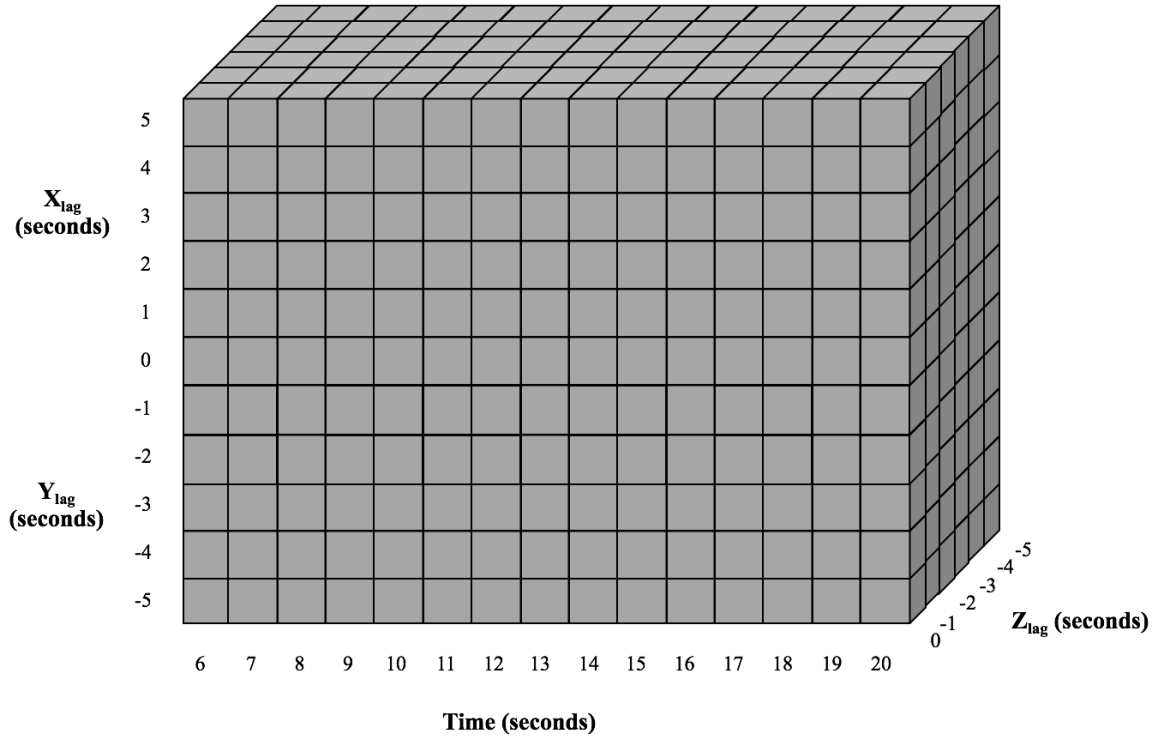

Figure 9. An example trivariate windowed cross-correlation (WCC-3) tensor. Here, each cube represents a total correlation value at a lag position in X, Y, and Z, and at a given reference time point. The matrix slice from Figure 8 is associated with reference time point 6 seconds, and is therefore the first matrix slice on the left.

The TASS Algorithm

The Threeway Approximate Spatiotemporal Symmetry (TASS) algorithm functions much the same way that the Pairwise Approximate Spatiotemporal Symmetry (PASS) does, with additional considerations for the three-signal nature of the analysis. The algorithm has the following steps:

1. The trivariate windowed cross-correlation (WCC-3);
2. The subsequent correlation matrix;
3. The multivariate distance; and
4. The symmetry indication.

In the first step, the association between three signals is captured using WCC-3. This process uses lagged windows to paint a complete picture of short-term dynamics in the relationship between three signals. The result is a tensor, where one dimension is reference time, another dimension is lag from one signal, and the last dimension is bipolar and contains lags from two signals. The second step correlates a matrix slice at a given reference time point with all successive matrix slices, so as to understand how the correlation pattern in a matrix slice changes across time. A subsequent correlation matrix is created, which contains only the most relevant of these time-based correlation patterns. In the third step, it must be determined which of these time-based correlation patterns exhibit similarity and which do not. Multivariate distance is used to do this. Those time-based correlation patterns which are near the center of the multivariate space exhibit similarity, and those far from the center do not. In the fourth and final step, a pre-determined threshold is used to distinguish those time-based correlation patterns which are near the center and which are far. In this way, we can distinguish time points which exhibit symmetry and which do not.

To capture the full extent of the symmetry relationships between three signals, it is necessary to consider relationships that evolve forward and backward in time. This is a necessity as a result of step 2, where the pattern of relationships is considered across time. If only forward-evolving, or *prospective* relationships are considered, only the beginning of a

symmetry segment is indicated; backward-evolving, or *retrospective*, relationships allow capture of the end of a symmetry segment. For this reason, the analysis will be run in both temporal directions. To achieve retrospective symmetry indication, both time series are reversed, and each of the following steps are computed on both the prospective and retrospective time series. The retrospective WCC-3 tensor is reversed again before step 2 so that retrospective relationships can be understood in step 2 and beyond as they unfold in prospective time. See Sjobeck et al. (n.d.) for a justification of this decision.

Subsequent Correlation Matrix. Symmetry is expected in both spatial and temporal dimensions. Association between sources at any given time constitutes the spatial symmetry component, and is measured by the WCC-3. The temporal component of symmetry is present when symmetry is maintained across time. It is therefore necessary to measure symmetry with particular reference to *how long symmetry is maintained across time*. The TASS algorithm accomplishes this by correlating matrices in the WCC-3 tensor associated with a given reference time point with matrices associated with successive reference time points. In this way it is possible to determine how the pattern of spatial symmetry changes with each new time point. If spatiotemporal symmetry is present, then it is expected that the pattern will remain invariant for some amount of time. The goal of this step is to determine how long on average the spatial symmetry pattern is expected to be maintained across time, and then create a matrix which reflects these time-based correlations of correlations. This matrix will be referred to as the *prospective correlation matrix* whenever forward-evolving relationships are considered, and *retrospective correlation matrix* whenever backward-evolving relationships are considered. Without loss of generality, we will use the prospective correlation matrix to exemplify the process by which both the prospective and retrospective matrices are produced.

In order to produce the prospective correlation matrix, it is necessary to consider what will be called the correlated windowed cross-correlation (CWCC) matrix. This matrix is an intermediary step between WCC-3 and the prospective matrix which captures the

correlational relationships between the spatial symmetry patterns at all time points. The CWCC matrix is created by correlating lagged relationships contained in a matrix slice of the WCC-3 matrix with lagged relationships contained in all other matrix slices of the WCC-3 matrix. The general form of the CWCC matrix is

$$CWCC = \begin{pmatrix} \mathbf{r_{\hat{t}_1\hat{t}_1}} & r_{\hat{t}_1\hat{t}_2} & r_{\hat{t}_1\hat{t}_3} & \cdots & r_{\hat{t}_1\hat{t}_j} \\ r_{\hat{t}_2\hat{t}_1} & \mathbf{r_{\hat{t}_2\hat{t}_2}} & r_{\hat{t}_2\hat{t}_3} & \cdots & r_{\hat{t}_2\hat{t}_j} \\ r_{\hat{t}_3\hat{t}_1} & r_{\hat{t}_3\hat{t}_2} & \mathbf{r_{\hat{t}_3\hat{t}_3}} & \cdots & r_{\hat{t}_3\hat{t}_j} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_{\hat{t}_i\hat{t}_1} & r_{\hat{t}_i\hat{t}_2} & r_{\hat{t}_i\hat{t}_3} & \cdots & \mathbf{r_{\hat{t}_i\hat{t}_j}} \end{pmatrix} \tag{21}$$

Notice that the diagonal values, shown in bold, are all equal to 1 as they indicate the correlation of a column of the WCC matrix, which represents a given reference time point, with itself.

The CWCC matrix contains all possible correlations between total correlation patterns associated with all time points, and therefore contains more information than is necessary for understanding symmetry relationships. To start, the CWCC matrix contains correlational information about time points that both precede and succeed each reference time point. Because of the way that prospective and retrospective analyses are structured, it is only necessary here to consider subsequent time points. In addition, it is obvious that, at some subsequent point in time, the relationship between the reference time point and its subsequent neighbors will cease to reflect potential symmetry relationships. Time points far from any given reference time point are unlikely to be highly correlated with the current time point, and if they are, then likely this is the result of spurious relationships. For this reason, only so many time points will need to be retained for further analysis steps. This number of time points will be called the *prospective value* or *retrospective value*.

Several criteria for choosing a prospective value are relevant. In general, the TASS algorithm selects a prospective value based on the number of time points which maintain

some relation to the reference time point. The chosen subsequent time point is the first subsequent time point to share no similarity with the reference time point. It is likely that the correlations between the reference time point and time points beyond the chosen subsequent time point will remain in a bounded region around zero and therefore contain little that is relevant to potential symmetry.

The most common case is the one illustrated above: at some point in time, the reference time point will no longer be related to subsequent time points. This corresponds to a correlation that is either zero or negative. Because windows are only incremented by one, the spatial symmetry pattern at each reference time point contains overlapping information with the spatial symmetry pattern of its neighboring time points, and this guarantees that the prospective correlation pattern will always begin as a decreasing function of time. The first subsequent time point which has a zero or negative correlation with the reference time point indicates the extreme end of a segment of time defined by how long the potential spatial symmetry pattern is maintained across time. This segment of time exemplifies the dwindling of relatedness that a time point has with its nearest neighboring time points, and thus contains the desired relational information bounded by a reasonable stopping point. In practice, the prospective value in this case is determined by the point at which the correlation first changes sign from positive to negative.

It is not necessary, however, that the similarity pattern at one time eventually stop exhibiting similarity with subsequent time points. Periodic data, for instance, have a prospective correlation pattern that never reaches zero. Instead, the prospective correlation pattern oscillates at a value higher than zero. To account for this possibility, another approach is taken. What is desired in this case is the first subsequent time point at which the correlation pattern gets as close as it can be to maximally unrelated before increasing in relatedness. In other words, there naturally exists a set of peaks and valleys in the correlation pattern, and the desired prospective value is the point at the very bottom of the first valley. The minimum is estimated as the point at which the first derivative of the

correlation pattern changes sign, from negative to positive. This current method uses generalized local linear approximations (GLLA) as outlined in Boker, Deboeck, Edler, and Keel (2010) to determine this change in first derivative. A time delay embedding matrix, formed from the vector of prospective correlations, can then be multiplied by a transformation matrix to estimate the first derivative of the vector. In this study, an embedding dimension of 4 was specified for each example where this method was needed.

A prospective value is determined for each reference time point in the CWCC matrix. In order for the prospective correlation matrix to be a rectangular matrix with a consistent number of prospective correlations in each row, it is necessary to aggregate across all prospective values to get one representative prospective value for each time point. This means that each row of the prospective correlation matrix may contain too few relevant or too many irrelevant prospective correlations, and consistutes an inherent limitation of the method. Aggregation is done using the median of scores. The median is chosen here because it is not subject to outliers, which are likely in the distribution of prospective values. The median prospective value informs the number of columns of the prospective correlation matrix. Additionally, the number of rows will be shortened to accommodate for the last row to contain the appropriate number of prospective correlations.

The prospective correlation matrix, therefore, contains only those subsequent correlation values which have the potential to indicate symmetry. The general form for the prospective correlation matrix, PCM, is

$$PCM = \begin{pmatrix} \mathbf{r_{\hat{t}_1\hat{t}_1}} & r_{\hat{t}_1\hat{t}_2} & r_{\hat{t}_1\hat{t}_3} & \cdots & r_{\hat{t}_1\hat{t}_p} \\ \mathbf{r_{\hat{t}_2\hat{t}_2}} & r_{\hat{t}_2\hat{t}_3} & r_{\hat{t}_2\hat{t}_4} & \cdots & r_{\hat{t}_2\hat{t}_p} \\ \mathbf{r_{\hat{t}_3\hat{t}_3}} & r_{\hat{t}_3\hat{t}_4} & r_{\hat{t}_3\hat{t}_5} & \cdots & r_{\hat{t}_3\hat{t}_p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{r_{\hat{t}_{i-p}\hat{t}_1}} & r_{\hat{t}_{i-p}\hat{t}_2} & r_{\hat{t}_{i-p}\hat{t}_3} & \cdots & r_{\hat{t}_{i-p}\hat{t}_p} \end{pmatrix} \tag{22}$$

where p is the prospective value, $\hat{t}_i$ is the first time point in the PM, $\hat{t}_{i+p}$ is the time point

in the PM that is p time points from $\hat{t}_i$, and r is the correlation between the lagged correlations in the windowed cross-correlation matrix associated with $\hat{t}_i$ and the lagged correlations in the windowed cross-correlation matrix associated with $\hat{t}_{i+p}$. Note that the time index is $\hat{t}_i$ instead of t_i; this is because the windowed cross-correlation process does not contain the same points as the original time series. The number of columns is dictated by the median prospective value, and this is necessary so that each row in the matrix is of the same length. The number of rows is also dictated by the median prospective value, but this is instead because any time point beyond the median subsequent value must have less than that many column values and therefore cannot be included.

Multivariate Distance. The raw prospective correlation matrix is informative but does not yet indicate which time points are symmetric and which are not. A method for deciding which time points are reflective of symmetry is needed. The TASS algorithm uses multivariate distance to accomplish this, particularly squared Mahalanobis distance (Mahalanobis, 1936). Squared Mahalanobis distance allows for multivariate data near the multivariate mean to be distinguished from multivariate data far from the multivariate mean. In this case, prospective correlation patterns which are near the mean of all prospective correlation patterns should be considered symmetric, while prospective correlation patterns which are far from the mean of all prospective correlation patterns should be considered non-symmetric.

The formula for squared Mahalanobis distance is as follows:

$$D_i^2 = (x_i - \mu)' \Sigma^{-1} (x_i - \mu) \tag{23}$$

where x_i is the ith row of the data matrix x, μ is a vector of the mean values of the columns of x, Σ is the covariance matrix of x, and D_i^2 is the squared Mahalanobis distance of the ith row of x. Each squared Mahalanobis distance value is a reflection of the relationship between a given time point and all other points, such that lower squared

Mahalanobis distance values indicate similarity near the expected similarity value of all time points and higher squared Mahalanobis distance values indicate similarity that is far removed from the expected similarity value of all time points.

Squared mahalanobis distance is often used in outlier detection (Ben-Gal, 2010; Leys, Delacre, Mora, Lakens, & Ley, 2019). Squared mahalanobis distance is assumed to follow a chi-square distribution, and so chi-squared will be used to threshold nearness and farness of squared Mahalanobis distance values. Because the prospective correlation matrix exists in two dimensions, a degrees of freedom value of 2 will be used in this case. The two dimensions here represent the spatial dimension and the temporal dimension, where along both simultaneously symmetry is expected to exist. A significance level of .001 will be used, as is consonant with standard procedures in outlier detection.

Time series which are highly symmetric will still be found to produce multivariate outliers with squared Mahalanobis distance, because squared Mahalanobis distance is sensitive to any deviations relative to the system. For this reason, it is necessary to establish a threshold for symmetry above which squared Mahalanobis distance indication is not needed. If it is that the similarity in a row of the prospective correlation matrix is maintained above this threshold, we can say with certainty that that reference time point is symmetric without using its squared Mahalanobis distance value to make the determination. The threshold can be set by the user; the current analysis uses a threshold of .90, which is to say that any reference time point which maintains a correlational similarity of .90 with all of the future time points in its row will be considered symmetric. For plotting purposes, time points which meet this criteria are assigned squared Mahalanobis distance values of 0.

Symmetry Indication. Squared Mahalanobis distance allows for time points indicating symmetry to be distinguished from time points which do not indicate symmetry. In this way, all time points which have been included in the analysis may be indicated as either symmetric or non-symmetric. Because prospective and retrospective analyses

provide two different perspectives as to what constitutes symmetry, it is also possible to indicate time points as either indicative of non-symmetry, prospective symmetry, retrospective symmetry, or both prospective and retrospective symmetry. This information can be shown visually, and exploratory information can be gleaned in this way.

Additionally, it is possible to aggregate this discrete symmetry information into a metric which can be used in confirmatory analyses. Several symmetry metrics are offered:

1. Symmetry ratio, i.e., proportion of the time series deemed symmetric;
2. Mean length of symmetry segments; and
3. Standard deviation of symmetry segments.

The symmetry ratio is a ratio of the number of time points associated with squared Mahalanobis distance values below a chosen threshold to the number of total time points in the system. The ratio can be written in the following form:

$$S = \frac{|D_t^2 < \hat{D}^2|}{|D_t^2|}, \tag{24}$$

where S represents the proportion of symmetry in the global system, D_t^2 is the set of squared Mahalanobis distance values associated with each time point t, and $\hat{D}^2$ is the threshold value. This symmetry value then represents the proportion of time points in the system which reflect symmetry greater than the value associated with the chosen threshold. The mean and standard deviation length of symmetry segments are calculated based on all segments in the time series which have been determined by the method to be symmetric. The mean or standard deviation is then taken across the lengths of each of these segments. If less than two segments appear in the data, then the mean and standard deviation values cannot be used.

Analyses

The aim of this study is to validate the TASS Algorithm, by analyzing representative data from both simulation and empirical research. Simulated data demonstrates how the

TASS algorithm performs under known symmetry conditions. Empirical data demonstrates the utility of the TASS algorithm in analyzing data collected in empirical environments, where the goal is to understand the relationship between symmetry and other variables of interest. It is important to note that this study will not exhaustively test any particular psychological theory of interpersonal or intrapersonal dynamics.

The following simulated examples were chosen to cover extreme cases from which the TASS algorithm might be asked to extract symmetry segments. Some are chosen because their symmetry pattern is clearly evident, and some are chosen to determine how the pairwise symmetry patterns differ from the threeway symmetry pattern. Each of the three signals in each simulation condition is 900 samples in length, and each simulation condition was replicated 500 times. Sine curves are used to understand how the TASS algorithm handles periodic data, where symmetry is expected to be high. Alternatively, simulated noise will be used to understand how the TASS algorithm responds to data with no discernible pattern. Representative time series for each simulation condition can be found in Figure 10.

To determine how the TASS algorithm performs when all three signals are perfectly periodic, the first simulation condition, Ex1, was comprised of three sine curves ($f = 25$ Hz). All time points in Ex1 were expected to be indicated as symmetric. To determine how the TASS algorithm performs when periodic signals are collected with some noise, the second simulation condition, Ex2, the third simulation condition, Ex3, and the fourth simulation condition, Ex4, were each comprised of three sine curves with increasingly more variable amounts of noise added (0.25, 0.50, and 1.00, respectively). In all three of these conditions, segments of symmetry are expected, but broken up by segments of non-symmetry in a random manner. As the noise variance increases, less symmetry is expected. To determine how the TASS Algorithm performs when two signals are related and one is unrelated, the fifth simulation condition, Ex5, was comprised of two sine curves and one noise signal. Here, complete symmetry indication is expected, as two related

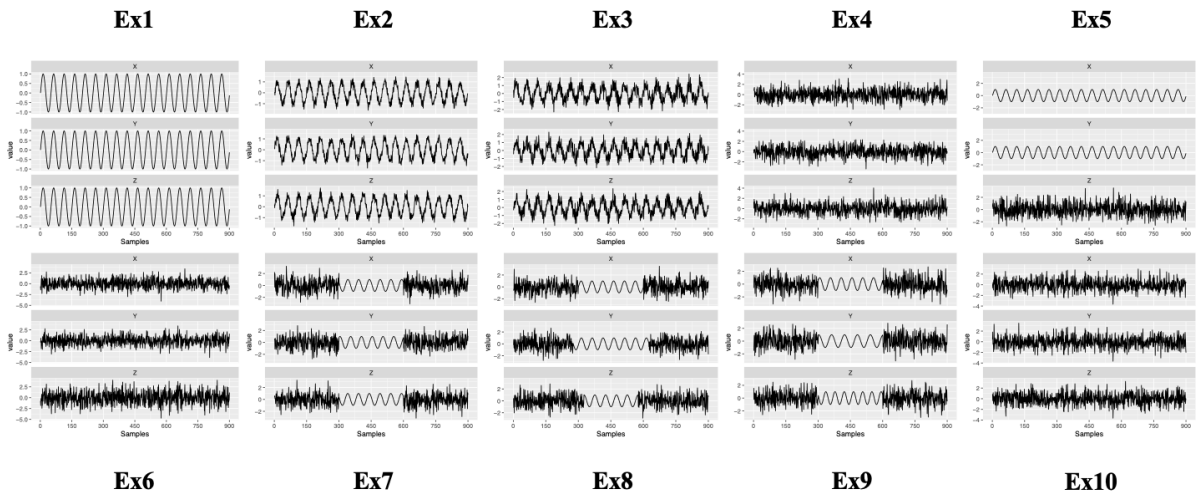

Figure 10. Representative time series for each simulation condition. Unless otherwise stated, sine curves have a frequency of 25 Hz and noise is sampled from a standard normal distribution, $N(0, 1)$. **Ex1**, three sine curves; **Ex2**, three noisy sine curves, N(0,.25); **Ex3**, three noisy sine curves, N(0,.50); **Ex4**, three noisy sine curves, N(0,1); **Ex5**, two sine curves and one signal of noise; **Ex6**, two noise signals and one composite noise signal; **Ex7**, three noise signals with a sine curve segment; **Ex8**, three noise signals with sine curve segments that start and end at different time points; **Ex9**, three noise signals with sine curves segments of different frequencies (f_X = 25 Hz, f_Y = 30 Hz, f_Z = 21.43 Hz); and **Ex10**, three noise signals with a common noise segment.

signals should be all that is required for the TASS algorithm to detect symmetry.

To determine how the TASS algorithm performs when two signals are unrelated but the third is a linear composite of the first two, the sixth simulation condition, Ex6, was comprised of two noise signals and one signal which was a sum of the two noise signals. Here, broken symmetry is expected, because the influence of the other noise signal should diminish the symmetry relationships between one noise signal and the composite noise signal. To determine how well the TASS algorithm can extract a segment of symmetry from non-symmetric time series, the seventh simulation condition, Ex7, the eighth

simulation condition, Ex8, and the ninth simulation condition, Ex9, are all comprised of sine curve segments in the middle of noise signals. In Ex7, all three sine curve segments have the same frequency, $f = 25Hz$, and start and end at the same time points. Here, symmetry is expected from the start to the end of the sine curve segment. In Ex8, all three sine curve segments have the same frequency, but start and end in different places. The sine curve segment in time series Y starts 25 time points before and ends 25 time points after the sine curve segment in time series X. Furthermore, the sine curve segment in time series Z starts 25 samples after and ends 25 samples before the sine curve segment in time series X. Here, symmetry is expected to exist for the length of the sine curve segment in time series X. In Ex9, all three sine curve segments begin and end at the same time points, but each sine curve segment has a different frequency ($f_X = 25$ Hz, $f_Y = 30$ Hz, $f_Z = 21.43$ Hz). Here, broken symmetry is expected from the start to the end of the sine curve segment, as historically the PASS algorithm has struggled with differences in frequency. Lastly, to determine how the TASS algorithm performs when the segment of symmetry is non-periodic, the tenth simulation condition, Ex10, was comprised of three noise signals with a common noise segment. Here, symmetry is expected for the full length of the common noise segment.

The TASS algorithm was used to analyze threeway symmetry relationships between signals of each of the ten simulation conditions. In order to understand how these threeway symmetry relationships relate to each of the three pairwise symmetry relationships, the TASS algorithm metrics will be compared to PASS algorithm metrics. These metrics include the proportion of symmetric time points and the length of each symmetry segment. Together the threeway symmetry relationships and the three pairwise relationships will provide clarity regarding which pair of signals are driving the symmetry relationships between all three signals. Additionally, to understand how similar the threeway and pairwise results are, a new metric is offered, the mean absolute difference between the pairwise and threeway squared Mahalanobis distance values, or the MAD SMD. Note that

each MAD SMD compares pairwise and threeway SMD values for prospective and retrospective analyses separately.

The first empirical dataset of interest was selected to elucidate differences in the symmetry pattern of connected brain regions during a content-relevant task. In order to reduce analysis time, a reduced set of data was analyzed. Participants in this study were asked to attend to changes in greyscale images of a face superimposed over a house. In six of the trials, participants were asked to determine if the face changed between successive images; in the other six trials, participants were asked to determine if the house changed. Here, only data from one person will be analyzed, across each kind of trial. The TASS algorithm was used to capture the threeway symmetry relationships in the fMRI signals of three brain regions: the dorsolateral prefrontal cortex, the fusiform gyrus, and the superior temporal sulcus. These brain regions have been recognized as important in facial recognition. Additionally, the PASS algorithm was used to capture pairwise relationships between each pair of brain regions. In order to demonstrate the effect of the window size on the symmetry pattern, each time series was analyzed with the following window sizes: 10, 20, 35, and 50. Because of the length of "attend to houses" and "attend to faces" trials, and the fact that these trials alternate, whole time series were indicated for symmetry, and then the symmetry results were divided by trial. More symmetry was expected to be present between these three regions when attending to faces than when attending to houses. A t-test was used to determine if the lengths of symmetry segments in the "attend to faces" condition significantly differ from the lengths of symmetry segments in the "attend to houses" condition.

The second empirical dataset was selected to elucidate differences in the symmetry pattern of triads participating in a social task. This data comes from a study on the symmetry between groups of three drummers. Each triad of drummers were asked to drum along to a tempo that was either predictable or unpredictable. To demonstrate the utility of the algorithm without completing a full analysis, only 5 randomly-selected triads from

each condition were chosen and analyzed. The TASS algorithm was used to capture the threeway symmetry relationships in the heart rate, as measured by the interbeat interval (IBI), of all three drummers, and the PASS algorithm was used to capture the pairwise relationships between each pair of drummers. It was expected that more symmetry would be present when the tempo is predictable than when the tempo is unpredictable. Pairwise symmetry relationships were expected to be similar to threeway symmetry relationships in the predictable tempo condition. For groups with an unpredictable tempo, little or no symmetry was expected in both the pairwise symmetry pattern and the threeway symmetry pattern. Additionally, a mixed-effects model was used to determine if the lengths of symmetry segments between triads who drummed to a predictable tempo are significantly different from those triads that drummed to an unpredictable tempo. It is acknowledged that these results may be underpowered, but are given here to demonstrate how a symmetry analysis might reasonably be done.

Results

Simulations

The WCC3 plots for each simulation condition can be found in Figure 11. Because the WCC3 tensor produces a cube of total correlation values, only a matrix slice of the WCC3 tensor will be plotted. Each colored square represents the total correlation between three windows, one in each time series. White represents no association, while red represents perfect association. Each plot has two quadrants; the lower half contains the total correlation values for windows at various lags in the X and Y time series, and the upper half contains the total correlation values for windows at various lags in the X and Z time series. In the lower half, the window in the Z time series is kept at the default position, where $\tau = 0$. Likewise, in the upper half, the window in the Y time series is kept in the default position. Each lag value represents a number of steps backwards in the time series. Each simulation condition used the same window size, $w_{size} = 50$ samples, and the same

maximum lag, $\tau_{max} = 50$ samples. Note that each analysis produces a prospective and retrospective WCC3 tensor, and so for simplicity only the prospective WCC3 plot will be given.

The WCC3 plot for Ex1 (i.e., three sine curves) suggest a lattice pattern of total correlation values. The moments of high total correlation occur when each of the three time series is either at $\tau = 0$, $\tau = 25$, or $\tau = 50$. This is because the frequency of all three in-phase sine curves is 25 Hz. The lines that connect these moments of high total correlation are when two signals exhibit association, as the other is being lagged to the next peak association. The WCC3 plots for Ex2, Ex3, and Ex4 reflect the noise that has been added to the sine curves in increasing amounts. In the WCC3 plot for Ex2, the lattice pattern seen in Ex1 has been blurred, and the distinction between red and white is less obviously present. In the WCC3 plot for Ex3, the lattice pattern is still evident, but the moments of high association have bled together, producing a heat map that is predominantly red and orange. The WCC3 plot for Ex4 captures the absence of all pattern, as too much noise has been added to the sine curves for their periodic nature to be evident.

The WCC3 plot for Ex5 (i.e., two sine curves and a noise signal) illustrate a common phenomena in WCC3 plots, vertical and diagonal bands of association. In the upper quadrant, the vertical bands suggest that a relationship is present between Y at $\tau_Y = 0$, X at $\tau_X = 0$, $\tau_X = 25$, and $\tau_X = 50$, and Z at all lags. Therefore, a relationship between X and Y exists at these specified lags regardless of the value of Z. This relationship occurs for several lags in X because of the periodic nature of X, a sine curve. Similarly, the diagonal bands in the lower quadrant suggest that a relationship is present between Z at $\tau_Z = 0$ and X and Y as their lags increase together, starting at intervals of $\tau_X = 25$. Ultimately, it is apparent that Z, the noise signal, has no bearing on the relationship between X and Y, the two sine curves.

The WCC3 plot for Ex6 (i.e., two noise signals and a linear composite of the two noise signals) reflects a very noisy association pattern. Here the strongest association is seen

when τ_Y and τ_Z are both 0, across all τ_X. This suggests a relationship between Y and Z at 0 lag, irrespective of the values in X. Additionally, there is association present in the lower quadrant, when Z is at $\tau_Z = 0$, X is at $\tau_X = 0$, and across all τ_Y. This suggests that X and Y are both related to Z but neither X nor Y are related to each other.

The WCC3 plot for Ex7 (i.e., three noise signals with sine curve segments) exhibits a lattice pattern, like that seen in the WCC3 plot for Ex1; however, because the sine curve segments have random phase, the moments of high association are not found at intervals of 25 sample lags. The WCC3 plot for Ex8, where the sine curve segments do not begin and end at the same time points, reflects this inconsistency. Note that while evidence of the displaced lattice pattern in Ex7 exists in the lower quadrant, more than half of the upper quadrant is missing the lattice pattern. This is likely because time series Z has a shorter sine curve segment that starts much later in time than time series X and Y, and so this WCC3 plot reflects the transition from noise to sine curve segment in Z. The WCC3 plot for Ex9, where the sine curve segments have different frequency, exhibit evidence of the bleeding of strong association seen in Ex3, suggesting that differences in frequency produce a similar effect to adding noise.

The WCC3 plot for Ex10 (i.e., three noise signals with a common noise segment) suggest a minimal relationship between the three signals during the common segment of noise, because the relationship between three of the same signal of noise is only present at equivalent lags. For this reason, the strongest association is present when all windows are at $\tau = 0$. Because the WCC3 tensor also captures bivariate relationships, there is also a strong association present across all τ_X when $\tau_Y = 0$ and $\tau_Z = 0$, suggesting a relationship between Y and Z at 0 lag. The diagonal line of association in the lower quadrant suggests a relationship between X and Y at all equivalent lags when $\tau_Z = 0$. Similarly, the diagonal line of association in the upper quadrant suggests a relationship between X and Z at all equivalent lags when $\tau_Y = 0$. Thus, at this moment in the WCC3 tensor, all three signals are related when at least two signals are at equivalent lags.

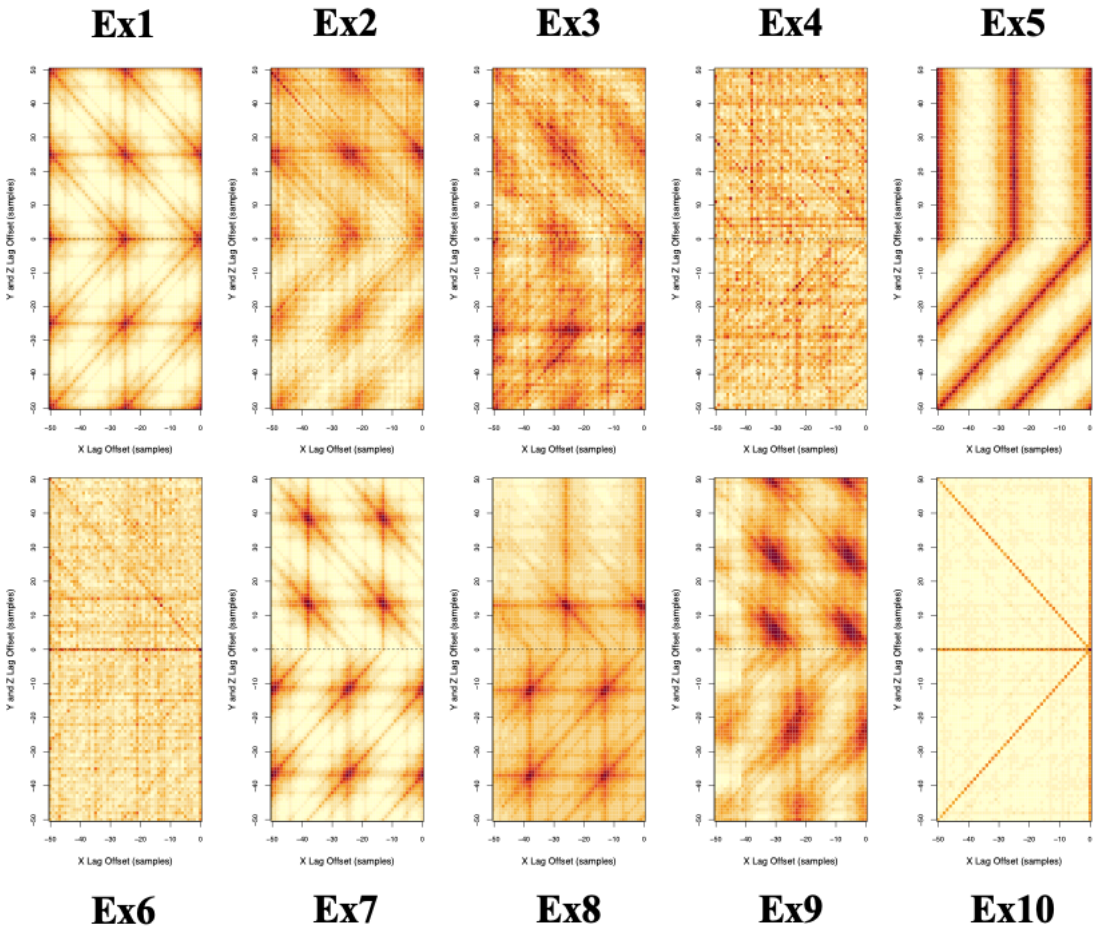

Figure 11. Prospective WCC3 plots for each simulation condition. The lower quadrant represents the lagged and windowed total correlation pattern between time series X and Y, while the upper quadrant represents the lagged and windowed total correlation pattern between time series X and Z. White pixels represent no association, while red pixels represent perfect association. For all conditions, w_{size} = 50 samples, τ_{max} = 50 samples.

Symmetry segmentation plots for the simulated examples can be found in Figure 12. Here, blue indicates a time point that the TASS algorithm has determined to be symmetric, grey indicates a time point that the TASS algorithm has not determined to be symmetric, and red indicates time points that are not included in the analysis. The symmetry segmentation plots for Ex1, Ex2, Ex3, and Ex4 demonstrate how much symmetry the TASS algorithm can extract from sine curves with increasing amounts of noise. The sine curves in Ex1 are indicated as perfectly symmetric, while the noisy sine curves in Ex2 and Ex3 exhibit broken symmetry, especially around the peaks and troughs. The sine curves in Ex4 have lost almost all of their periodicity, although the algorithm has still indicated that several time points are symmetric. The time series in Ex5 are also perfectly indicated as symmetric, suggesting that only two of three signals need to exhibit association in order for the TASS Algorithm to indicate time points as symmetric.

The symmetry segmentation plot for Ex6 demonstrates that the TASS Algorithm can detect the symmetry between a signal of noise and a linear combination of that signal with another signal of noise, although the symmetry here is also broken, likely because of the influence of the noise from the other signal. The symmetry segmentation plots for Ex7, Ex8, and Ex9 illustrate the effects that differences in frequency and starting and ending points have on the algorithm's ability to extract segments of sine curves from noise signals. In Ex7, where the segments start and end at the same time points and have the same frequency, the TASS Algorithm can extract segments of symmetry associated with the sine curves, although it underestimated its length evenly at the transitions. In Ex8, where the segments have different starting and ending time points, a segment of symmetry was extracted that is roughly the length of the middle segment length in X. In Ex9, where the segments have different frequencies, the symmetry segment boundaries are well determined, but the segment itself is fragmented, and symmetry artifacts are present in time points located in the noise segments. The symmetry segmentation plot for Ex10 demonstrates that the TASS Algorithm can extract a segment of noise common to three noise signals.

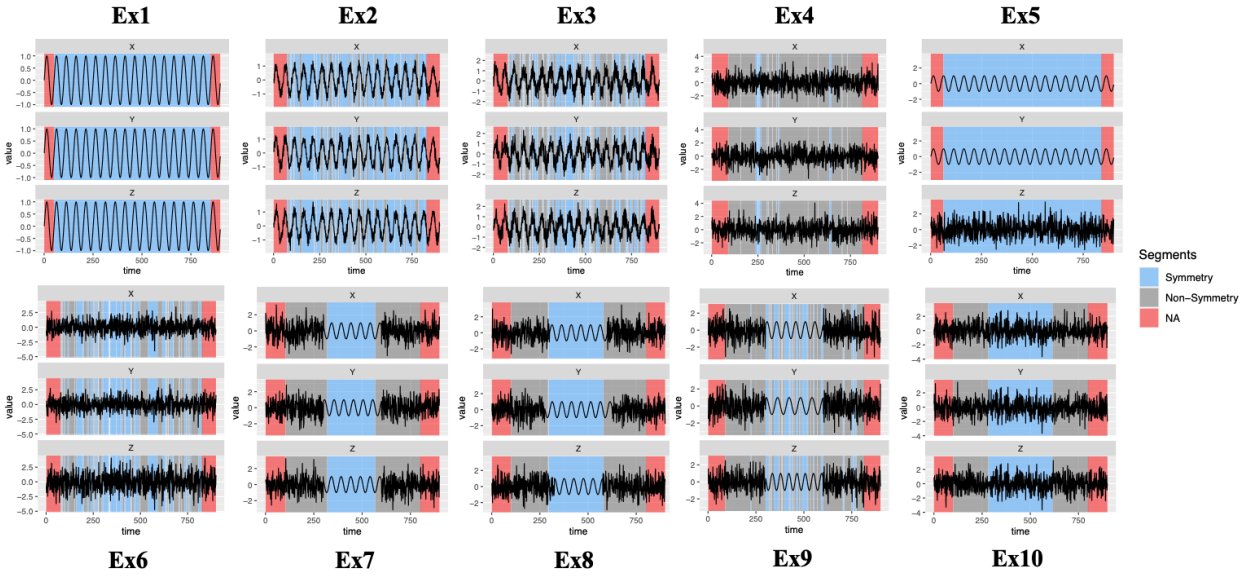

Figure 12. Symmetry segmentation plots for each simulated condition. Blue segments indicate symmetry, grey segments indicate non-symmetry, and red segments indicate time points not included in the analysis.

The proportion of symmetry was calculated for each simulation condition, aggregated across all 500 replications, to produce one representative average proportion of symmetry for that simulation condition. In addition to the threeway proportion of symmetry, pairwise symmetry proportions were also calculated, so that each pairwise relationship could be compared to the threeway relationship. All pairwise analyses also used $w_{size} = 50$ Hz and $\tau_{max} = 50$ Hz. It is worth noting, however, that the TASS algorithm uses total correlation, an information theory metric, while the PASS algorithm uses correlation as its association metric, and that some differences between threeway and pairwise relationships may be a result of this difference. All average symmetry proportions can be found in Table 1.

As expected, the average proportion of symmetry in Ex1 indicates that each of the 500 replications produced perfect symmetry indication, and this was also the case for each pairwise proportion. The threeway symmetry proportion drops as more noise is added in Ex2, Ex3, and Ex4, resulting in almost no symmetry in Ex4. The pairwise relationships for

Ex2 and Ex3, however, remain at perfect symmetry indication and do not decrease with the increase in noise. This is likely a result of improper choice of hyperparameters for the pairwise analyses. By Ex4, the pairwise relationships have completely disappeared, as indicated by a symmetry proportion of 0. The proportion of symmetry for Ex5 demonstrates perfect symmetry indication for the threeway case, and also between time series X and Y (in this case, the two sine curves), but only one-third of the time points between the pairs X and Z, and Y and Z, where Z is the noise signal. The proportion of symmetry for Ex6 suggests no symmetry between time series X and Y, the two noise signals, but that both exhibit some symmetry with time series Z, the linear combination of both noise signals. The threeway proportion indicates that on average half of the time points in the linear combination noise signal are related to either individual noise signal. The threeway proportions of symmetry for Ex7 and Ex8 indicate that about one-third of the time points are symmetric, and that corresponds to the proportion of time points in the time series that are sine curves. Like in Ex2 and Ex3, the pairwise relationships contain more symmetry time points than would be expected based on the length of the sine curve segment. In Ex9, where frequencies differ, a smaller threeway proportion of symmetry was found, likely due to broken segments of symmetry. The pairwise symmetry proportions seem to have done a better job of handling the frequency differences, indicating a more reasonable proportion of symmetry. The Ex10 symmetry proportions are consistent across threeway and pairwise analyses, and are consonant with expectations based on the length of the common noise curve segment.

Symmetry segment lengths for each simulation example, aggregated across all 500 replications, and calculated for both threeway and pairwise analyses, can be found in Table 2. Here, symmetry segment lengths are measured in samples. These results mirror those reflected by the symmetry proportions. Ex1 demonstrates full lengths of symmetry indication for both threeway and pairwise segment lengths. Ex2 and Ex3 have average pairwise segment lengths that have not dropped as significantly from Ex1 as the threeway

Condition	X & Y & Z	X & Y	X & Z	Y & Z
Ex1	1.00	1.00	1.00	1.00
Ex2	0.73	1.00	1.00	1.00
Ex3	0.64	1.00	1.00	1.00
Ex4	0.04	0.00	0.00	0.00
Ex5	1.00	1.00	0.33	0.33
Ex6	0.54	0.00	0.30	0.30
Ex7	0.38	0.99	0.99	0.99
Ex8	0.39	0.97	0.96	0.95
Ex9	0.22	0.43	0.40	0.45
Ex10	0.46	0.47	0.47	0.46

Table 1

The proportion of symmetry for each simulated example, aggregated across all 500 replications. Additionally, symmetry proportions are given for the threeway analysis and each of the three pairwise relationships. w_{size} = 50 Hz and τ_{max} = 50 Hz for all pairwise analyses.

pairwise segment lengths have. In Ex4, where much of the periodic nature of the sine curves has been obscured, a small average threeway segment length and no bivariate segment length. Ex5 demonstrates long threeway segment lengths, and this seems to be driven by the relationship between X and Y, which also has long symmetry segment lengths. No symmetry segment lengths are present between X and Y in Ex6, where X and Y are two signals of noise, but X and Y both have symmetry segments with Z. These symmetry segment lengths, however, are less than half of the average symmetry segment lengths between all three signals. The pairwise and threeway analyses for Ex7 and Ex8 indicate long symmetry segments, which are consonant with our expectations. The pairwise symmetry segment lengths for Ex8 indicate that Y, the time series with the

longest symmetry segment, and X, the time series with the second longest symmetry segment, have the longest pairwise symmetry segment length. For Ex9, where the sine curve segments have different frequencies, the TASS algorithm produces much shorter segments than the pairwise segments produced by the PASS algorithm. Additionally, the longest pairwise segment lengths existed between the sine curve segments with the shortest and longest frequency, which is counter to expectations, and the shortest pairwise segment lengths existed between the shortest and middle frequency. Ex10, the common noise segment example, demonstrates comparable symmetry segment lengths for the threeway and pairwise relationships, where all are roughly the length of the common noise segment.

Condition	$X\&Y\&Z$	$X\&Y$	$X\&Z$	$Y\&Z$
Ex1	801.00	801.00	801.00	801.00
Ex2	14.45	780.78	780.64	780.66
Ex3	10.64	776.39	776.35	776.33
Ex4	2.35	0.00	0.00	0.00
Ex5	775.81	801.00	5.95	5.94
Ex6	9.53	0.00	4.67	4.70
Ex7	163.27	246.16	230.59	251.26
Ex8	142.96	185.91	153.61	139.56
Ex9	6.43	89.58	41.07	174.67
Ex10	211.22	261.83	269.36	267.24

Table 2

Symmetry segment lengths for each simulation example, aggregated across all 500 replications, and calculated for both threeway and pairwise analyses. Each length is measured in samples.

The mean average difference (MAD) in squared Mahalanobis distance (SMD) values for each pair of time series can be found in Table 3. Here, each MAD SMD value represents

the average absolute difference between SMD values in the threeway analyses and SMD values in the pairwise analyses, averaged across all 500 replications. Lower values indicate less difference, and higher values indicate more difference. Ex1 suggests that no difference is present between the threeway and pairwise SMD values, for any pair and for both prospective and retrospective directions. This is because SMD values for perfectly related time series are manually set to zero, but this still reflects expected MAD SMD values. The MAD SMD values for Ex2, Ex3, and Ex4 are all close to 20, and consistent across pairs. In Ex5, the MAD SMD values between X and Y, the two sine curves, are nearly 0, while the MAD SMD values for X and Z and Y and Z are almost 40, indicating less similarity in values. For Ex6, the MAD SMD value for X and Y, the two noise signals, is higher than the MAD SMD values for X and Z and Y and Z, suggesting that the latter two pairs are more similar than the former pair. For Ex7, Ex8, and Ex10, all MAD SMD values are high, but prospective MAD SMD values are higher than retrospective MAD SMD values. For Ex9, the MAD SMD values between X and Y and X and Z are much higher than the MAD SMD values between Y and Z.

Facial Recognition Study

The bold signals from the single participant in this study can be found in Figure 13. Brown dotted vertical lines separate the time series into intervals associated with the twelve trials. The first trial was an "attend faces" trial, and each successive trial alternates between "attend faces" and "attend houses" conditions. Trivariate windowed cross-correlation (WCC3) plots and symmetry segmentation plots for each given window size can be found in Figure 14. By design, the information provided by the WCC3 plot becomes more granular as the window size and lag size increases. Additionally, it is evident that, as the size of the window increases, the amount of symmetry detected by the TASS algorithm decreases.

When $w_{size} = 10$, the WCC3 plot indicates a clear picture of the symmetry

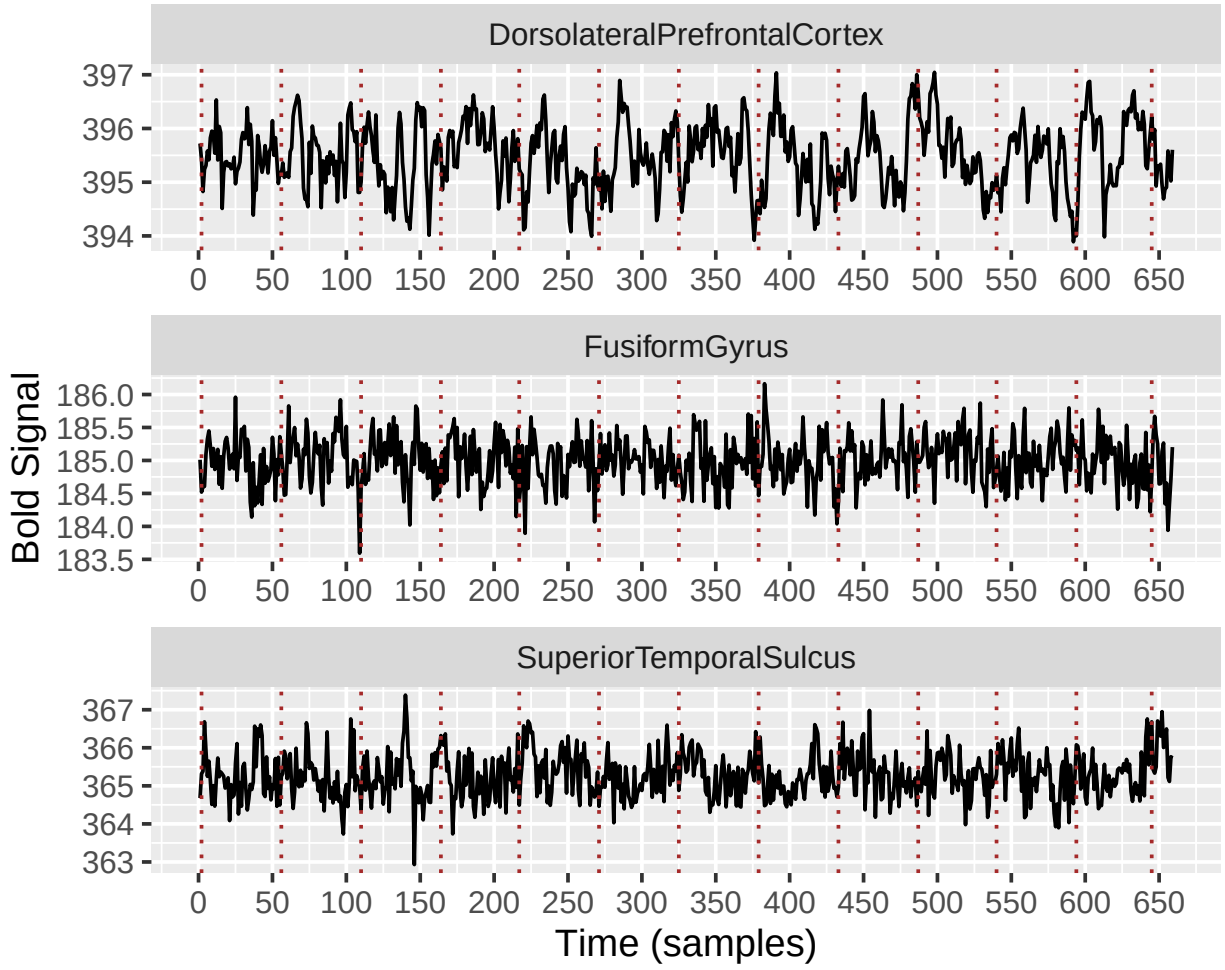

Figure 13. Bold signals from three brain regions (Dorsolateral Prefrontal Cortex, Fusiform Gyrus, and Superior Temporal Sulcus) from a single participant. Brown dotted vertical lines indicate the separation between the twelve different trials. The first trial is an "attend faces" trial, and each successive trial alternates between "attend faces" and "attend houses" conditions.

Condition	X & Y		X & Z		Y & Z	
	Pro	Retro	Pro	Retro	Pro	Retro
Ex1	0.00	0.00	0.00	0.00	0.00	0.00
Ex2	21.04	21.28	21.04	21.28	21.04	21.28
Ex3	22.58	22.70	22.58	22.70	22.58	22.70
Ex4	19.41	19.34	19.51	19.36	19.21	18.96
Ex5	0.02	0.01	39.15	37.71	39.14	37.71
Ex6	28.57	28.27	16.44	16.33	15.83	15.71
Ex7	52.08	38.07	52.84	38.13	52.61	38.46
Ex8	55.88	33.84	50.06	32.59	52.06	31.49
Ex9	36.11	35.07	35.40	34.79	24.67	23.45
Ex10	40.43	31.91	40.48	32.00	40.87	32.42

Table 3

The mean average difference (MAD) in squared Mahalanobis distance (SMD) values for each pair of time series in each simulation condition, averaged across all 500 replications. Lower values indicate less difference, and higher values indicate more difference.

relationships. A horizontal bar in the upper quadrant suggests that Fusiform Gyrus (FG) at $\tau_{FG} = 0$ and Superior Temporal Sulcus (STS) at $\tau_{STS} = 2$ are related regardless of the lag in Dorsolateral Prefrontal Cortex (DPC). The vertical bar in the lower quadrant suggests that DPC at $\tau_{DPC} = 6$ and STS at $\tau_{STS} = 2$ are related regardless of the lag in FG. The symmetry segmentation plot depicts full indication of symmetry. When w_{size} was increased to 20, several bars of high association appear in the WCC3 plot, suggesting a less clear relationship than was present when $w_{size} = 10$. Perhaps most illuminating in this plot are the isolated points of high association, where high association occurred at a single lag value in all three time series. The lower quadrant has one such point when $\tau_{DPC} = 16$, τ_{FG}

= 7, and $\tau_{STS} = 0$. The upper quadrant has another at $\tau_{DPC} = 6$, $\tau_{FG} = 0$, and $\tau_{STS} = 8$. The symmetry segmentation plot for $w_{size} = 20$ indicates segments of symmetry that often align with the beginning and ending of trials. Often the non-symmetric time points that separate segments of symmetry occur at trial transition points.

When $w_{size} = 35$, even less symmetry is apparent in the symmetry segmentation plots. At this window size, it is more evident that there are differences in the amount of symmetry between different trials. The WCC3 plot demonstrates a more complicated but still distinct difference between white and red pixels. It is also worth noting that, perhaps moreso than the WCC3 plots for $w_{size} = 10$ and $w_{size} = 20$, a lattice pattern is present. However, unlike the perfect symmetry simulation example, this plot is lacking in obvious moments of high association. This lattice pattern is maintained when w_{size} is increased to 50; however, there is much less distinction between red and white. Instead, the plot is dominated by less extreme and more homogeneous moments of association. This pattern translates into the least amount of symmetry present in symmetry segmentation plots when compared with smaller window sizes. Additionally, all time points of symmetry seem localized to two successive trials.

Symmetry proportions for the single participant can be found in Table 4, broken down by trial, window size, and whether the analysis was threeway or pairwise. A number of relationships are evident. First, consonant with the symmetry segmentation plots, increases in w_{size} translate to decreases in the symmetry proportion, for both houses and faces trials. This is the case for both threeway and pairwise analyses. When $w_{size} = 10$ and when $w_{size} = 50$, face trials and house trials have roughly the same amount of symmetry proportion. When $w_{size} = 20$, face trials exhibit more symmetry than house trials. However, when $w_{size} = 35$, house trials exhibit more symmetry than face trials. Consistently, threeway proportions are larger than pairwise proportions, regardless of the trial or the window size. The pairwise analyses for DPC and STS regularly produce the highest pairwise symmetry proportion.

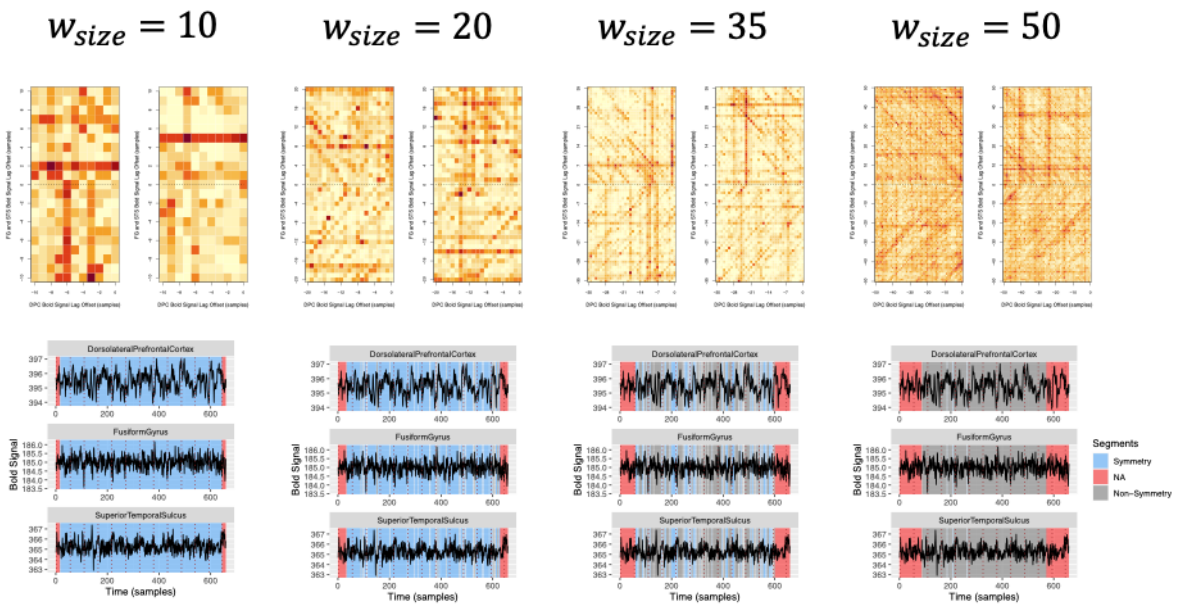

Figure 14. WCC3 and symmetry segmentation plots for one participant, across four different window sizes: 10, 20, 35, and 50. Window sizes and maximum lags are restricted to be the same. DPC = Dorsolateral Prefrontal Cortex, FG = Fusiform Gyrus, and STS = Superior Temporal Sulcus.

Table 5 contains the average symmetry segment lengths for house and face trials across all four window sizes, for the threeway and each pairwise analysis. Like with the proportions, the average symmetry segment length decreases as the window size increases, for both threeway and pairwise analyses. Additionally, the threeway symmetry segment lengths are on average longer than the pairwise symmetry segment lengths for both house and face trials and across all window sizes. For face trails, the pairwise analyses between DPC and FG demonstrate the longest average pairwise symmetry segment lengths, while no such consistent relationship exists for the house trials. Threeway symmetry segment lengths are comparable for house and face trials when $w_{size} = 10$, $w_{size} = 35$, and $w_{size} = 50$. When $w_{size} = 20$, face trials exhibit longer average threeway symmetry segment lengths

Trial	wSize	*DPC & FG & STS*	*DPC & FG*	*DPC & STS*	*FG & STS*
Faces	10	1.00	1.00	0.98	0.97
	20	0.94	0.60	0.67	0.67
	35	0.34	0.05	0.08	0.01
	50	0.01	0.00	0.00	0.00
Houses	10	1.00	0.99	0.98	0.98
	20	0.87	0.57	0.63	0.51
	35	0.38	0.12	0.13	0.01
	50	0.02	0.00	0.01	0.00

Table 4

Symmetry proportions between three regions of one participant's brain, as produced by the PASS and TASS algorithms. Symmetry proportions are broken down by trial type, window size, and whether or not the analysis was threeway or pairwise.

than house trials. When comparing the lengths of symmetry between house trials and face trials for this participant, no differences were found when $w_{size} = 20$ ($t(39) = $ -1.34, p $>$ .05) or when $w_{size} = 35$ ($t(53) = 0.60$, p $>$ .05).

Mean absolute difference (MAD) values for pairwise and threeway squared Mahalanobis distances (SMD) can be found in Table 6. On the whole, MAD SMD values increase as the window size increases, for both face and house trials. For pairwise analyses between DPC and FG, and at higher window sizes, face trials demonstrate less prospective MAD SMD and more retrospective MAD SMD than house trials. Additionally, for pairwise analyses between DPC and STS, and at higher window sizes, face trials demonstrate less prospective MAD SMD and comparable retrospective MAD SMD to house trials. For pairwise analyses between FG and STS, prospective and retrospective MAD SMD values are comparable.

Trial	wSize	*DPC & FG & STS*	*DPC & FG*	*DPC & STS*	*FG & STS*
Faces	10	41.46	35.09	26.84	18.28
	20	12.96	5.35	4.59	4.60
	35	2.66	2.08	1.87	0.80
	50	1.60	0.00	0.00	0.00
Houses	10	42.40	24.88	24.64	27.38
	20	8.77	4.51	6.43	3.87
	35	3.08	1.97	2.64	1.60
	50	2.00	0.00	0.80	0.00

Table 5

Average symmetry segment lengths for threeway and pairwise analyses broken down by face and house trials, and window sizes 10, 20, 35, and 50. symmetry segment lengths are represented in samples.

Drumming Study

Triads of drummers were asked to drum to either a predictable or unpredictable rhythm, and their heart rates were recorded. Data from two such triads, one with a predictable rhythm (the "synchronous" group) and one with an unpredictable rhythm (the "asynchronous" group), are plotted in Figure 15. These plots include time series plots, trivariate windowed cross correlation (WCC3) plots, and symmetry segmentation plots. The time series plots depict the periodic and oscillatory nature of these interbeat interval (IBI) signals. A standard window size and maximum lag of 20 were used in these analyses, for both Sync and Async triads. The prospective WCC3 plot for the Async triad has vertical bands in the upper quadrant, especially at $\tau_{122} = 2$ and $\tau_{122} = 15$. The fact that two such bands exist might reflect the periodicity of these signals. The vertical bands suggest a relationship between drummer 122 and drummer 123 irrespective of drummer

Trial	wSize	*DPC & FG*		*DPC & STS*		*FG & STS*	
		Pro	Retro	Pro	Retro	Pro	Retro
Faces	10	5.05	4.92	5.56	4.27	5.72	6.29
	20	10.63	9.27	9.72	8.68	10.25	9.35
	35	15.27	13.19	12.46	19.44	20.43	16.04
	50	15.50	23.05	17.83	18.88	20.26	21.55
Houses	10	4.75	5.10	4.64	5.43	5.32	5.28
	20	9.58	8.96	9.78	9.49	9.98	8.84
	35	18.69	15.20	19.98	13.11	15.76	17.79
	50	24.16	13.80	28.86	17.36	23.49	23.11

Table 6

Mean absolute difference values for threeway and pairwise squared Mahalanobis distances, across trials and window sizes.

124. Diagonal bands appear in the lower quadrant, further suggesting that drummer 122 and drummer 123 have a relationship when drummer 124 is at zero lag. Alternatively, the Sync triad has a prospective WCC3 which suggests a high association between all three signals when $\tau_{101} = 17$, $\tau_{102} = 0$, and $\tau_{103} = 1$. This suggest a close lag relationship between drummers 102 and 103, and perhaps that drummer 101 is leading the phenomenon. A horizontal bar exists at $\tau 103 = 1$, further suggesting a relationship between drummer 102 and 103 irrespective of the value of drummer 101, as well as a horizontal bar in both quadrants at $\tau_{101} = 17$, suggesting that drummer 101 has individual relationships with both drummers 102 and 103. While the WCC3 plot shows greater unity for the Sync triad, the symmetry segmentation plots suggest an over-indication of symmetry for both the Sync and Async triad, indicating no clear difference in symmetry segmentation. It is still worth noting, however, that the non-symmetric time points often

occur at the peaks in both Async and Sync triad time series.

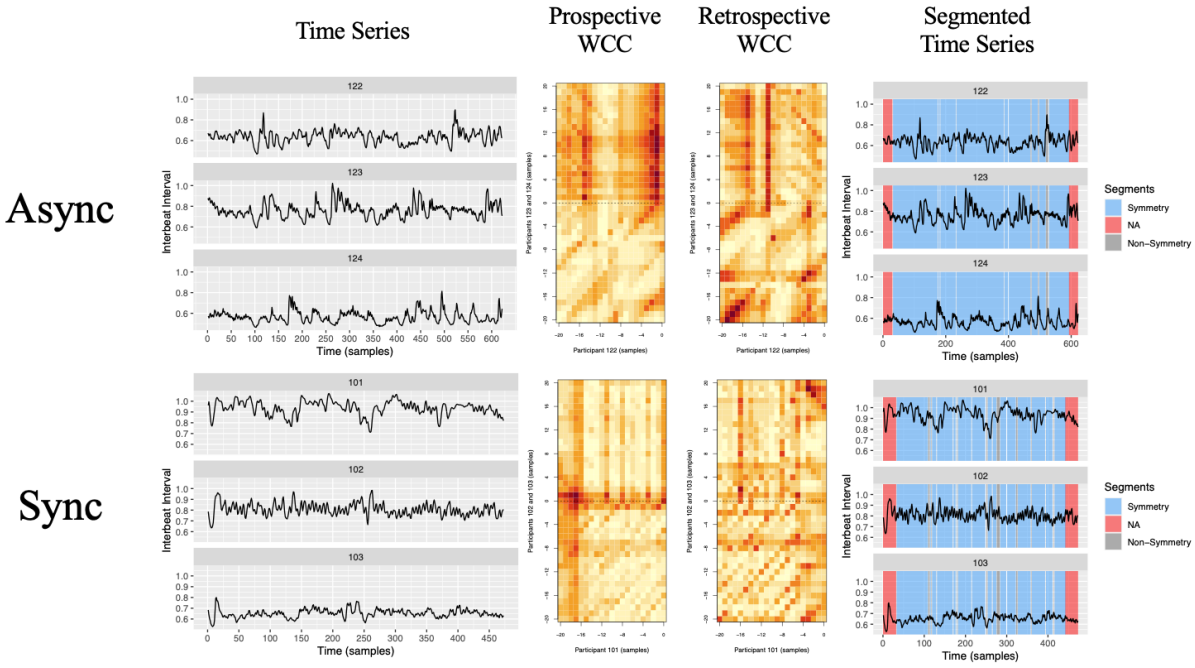

Figure 15. The time series plots, prospective and retrospective WCC3 plots, and the segmented time series plots for the chosen Async and Sync triads. $w_{size} = 20$, $\tau_{max} = 20$

Table 7 contains the proportions of symmetry for the threeway and pairwise analyses of each drumming triad in both the Async and Sync conditions. All threeway symmetry proportions seem much larger than would be expected, and were especially larger than the pairwise symmetry proportions. Additionally, no apparent differences exist between the Async and Sync triads. However, it is worth noting that the largest threeway symmetry proportion exists for a Sync triad, and the smallest threeway symmetry proportion exists for an Async triad. Further analyses would be needed to model differences in proportion between Async and Sync triads.

Table 8 contains the average symmetry segment length for each triad in each drumming condition, broken down by each threeway or pairwise analysis. Again, no clear distinction is present between Async and Sync triads in their average symmetry segment length. As

Condition	Triad	X & Y & Z	X & Y	X & Z	Y & Z
Async	1008	0.96	0.77	0.81	0.85
	1012	0.93	0.60	0.80	0.74
	1020	0.94	0.75	0.76	0.72
	1036	0.97	0.79	0.73	0.51
	1050	0.87	0.70	0.44	0.73
Sync	1001	0.91	0.81	0.79	0.71
	1004	0.95	0.71	0.71	0.78
	1006	0.94	0.75	0.66	0.76
	1011	0.99	0.74	0.79	0.90
	1013	0.97	0.68	0.79	0.79

Table 7

The proportion of symmetry for threeway and pairwise analyses of drumming triads in both the Async and Sync groups. X, Y, and Z were assigned to arbitrarily chosen members of the triad.

with the proportions, the longest average threeway symmetry segment length was from a Sync triad, and the shortest average threeway symmetry segment length was from an Async triad. Average threeway symmetry segment lengths were always longer than average pairwise symmetry segment lengths. Results from a multilevel model, where Sync or Async membership was used to predict symmetry segment lengths nested in triads, produced inconclusive results as to whether synchronous triads produced longer symmetry segment lengths than asynchronous triads, $\beta = 10.20$, $p > .05$. Mean absolute differences (MAD) in prospective and retrospective squared Mahalanobis distance (SMD) values, as found in Table 9, suggest that Async triads generally have greater MAD SMD values than Sync triads.

Condition	Triad	$X\&Y\&Z$	$X\&Y$	$X\&Z$	$Y\&Z$
Async	1008	27.20	6.56	9.61	8.67
	1012	10.89	2.67	6.95	6.38
	1020	11.02	5.64	7.10	4.97
	1036	28.22	7.43	5.27	3.34
	1050	7.22	4.92	2.81	5.65
Sync	1001	10.42	7.00	7.23	4.41
	1004	15.96	4.33	4.59	5.23
	1006	16.33	5.63	3.91	5.54
	1011	49.00	4.69	6.48	13.19
	1013	24.00	5.86	9.43	6.27

Table 8

The average symmetry segment length for each triad in each drumming condition, broken down by each threeway or pairwise analysis. Lengths are measured in samples.

Discussion

The TASS algorithm is a method that can be used to capture symmetry between three signals. Particularly, the TASS algorithm focuses on symmetry segmentation, and understanding how symmetry between signals is maintained across time. This is in contrast to standard time series methodology, which focuses on the magnitude or lag of association between signals. Likewise, standard methodology often assumes stationarity, i.e., that the properties of the time series remain consistent regardless of the interval or scale of the time series. The PASS algorithm, a method which precedes the TASS algorithm, captures bivariate relationships in a non-stationary way using product-moment correlation as an association metric. The TASS algorithm is an extension of the PASS algorithm that can be used to understand trivariate relationships. Instead of

Condition	Triad	*X* & *Y*		*X* & *Z*		*Y* & *Z*	
		Pro	Retro	Pro	Retro	Pro	Retro
Async	1008	14.09	13.82	11.67	13.12	10.33	10.63
	1012	12.51	12.16	8.09	8.71	9.89	10.33
	1020	10.20	11.01	10.24	11.60	10.09	10.13
	1036	10.51	9.11	12.42	11.78	14.62	14.65
	1050	12.33	11.77	14.00	13.19	11.00	11.17
Sync	1001	9.79	12.49	9.74	9.06	10.70	12.07
	1004	8.50	9.92	9.92	9.65	10.54	9.26
	1006	11.32	11.06	12.90	12.43	11.18	11.55
	1011	11.79	13.51	11.60	11.74	8.71	8.59
	1013	10.78	11.62	9.70	11.65	9.69	12.23

Table 9

Mean absolute difference values for prospective and retrospective squared Mahalanobis distance values, for each triad in each condition, Sync and Async.

product-moment correlation, which is inherently bivariate, the TASS algorithm uses total correlation, an information theory association metric that can be used to capture the association between any number of sets of data. The basis for the PASS algorithm, windowed cross-correlation, is similarly bivariate in nature; for this reason, the TASS algorithm relies on the newly-developed trivariate windowed cross-correlation (WCC3) procedure. The TASS algorithm provides individual information about symmetry segmentation that can be interpreted in a qualitative way, as well as symmetry metrics like proportion and segment length that can be used in model-based explorations.

In this study, the TASS algorithm was applied to a number of simulation conditions that were chosen to demonstrate how the TASS algorithm handles extreme circumstances.

These extreme circumstances are associated with known or expected symmetry relationships, and can provide information about whether or not the TASS algorithm performs in a way that matches a set of expectations. Consistently, the TASS algorithm did in fact conform to our expectations. The TASS algorithm was able to detect symmetry in periodic signals, found little or no association in the presence of noise, and captured segments of symmetry that were both periodic and aperiodic in a noisy signal. Noise that was added to perfectly periodic signals produced intermittent symmetry indication that could be informative of the breaking behavior of symmetry in applied data. It was found that time points in only two signals need be symmetric in order for the TASS algorithm to indicate those time points as symmetric. An aggregate signal, the sum of both other signals, may be related to both signals depending on how similar those signals are. Differences in frequency in periodic data are difficult for the TASS algorithm to detect as contiguously symmetric. All of these results indicate how well the TASS algorithm conforms to our expectations.

Additionally, the diversity in the association patterns found in the WCC3 tensors of these simulated signals provide a wealth of information that is interpretable in and of itself. For instance, lattice patterns were found to emerge for periodic data, and that the clarity of that lattice pattern changes in the presence of noise. Horizontal and vertical bands seem to indicate that one of the three signals is not providing association to the relationship of the other two signals, and diagonal bands suggest that at least two signals are associated across several lags. Perfectly associated signals that are aperiodic demonstrate a minimalist pattern in their WCC3 tensor. It also appears that strong association across too many lags is likely to translate to less symmetry indication, and that more symmetry indication occurs when the distinction between high and low association is clear. While the interpretations for WCC3 tensors and symmetry segmentation may coincide, it is apparent that WCC3 tensors provide information that may be used in answering other related questions about symmetry.

TASS algorithm symmetry metrics, particularly proportion and segment length, also conformed to our expectations. A smaller symmetry proportion and shorter symmetry segment lengths were present between perfectly periodic signals when more noise was added to them. Only two perfectly periodic signals were necessary for the algorithm to determine the entire time series was symmetric. Two noise signals were related to a linear composite a little more than half the time, in intermittent segments. The TASS algorithm extracted full segments of symmetry from signals of noise with sine curves in the middle, except for when the sine curves had different frequencies, in which case the symmetry proportion was less and the symmetry segments were shorter. Mean average differences (MAD) for squared Mahalanobis distance (SMD) values suggest that time series pairs with high similarity have lower MAD SMD values than time series pairs with low similarity. These results seem more influenced by the PASS algorithm results than the TASS algorithm results.

Data from two empirical studies, a facial recognition study and a drumming study, were used to get a sense for the ecological validity of the TASS algorithm. The goal for the facial recognition study was to determine if three regions of the brain relevant to facial recognition had bold signals that exhibited more symmetry when attending to changes in a face than when when attending to changes in a non-face object. If such evidence could be found, then these results could lend credence to these regions exhibiting functional connectivity, in that they exhibit levels of activation which occur in concert with each other during a facial recognition task. The bold signals from these regions were analyzed using four different window sizes: 10, 20, 35, and 50. TASS results suggested that symmetry was found between these signals, in varying amounts depending on the window size. When the window size was small ($w_{size} = 10$), the entirety of the time interval was determined to be symmetric. When the window size was large ($w_{size} = 50$), almost no symmetry was found. When the window size was 20 and 35, intermittent symmetry was present, and the pattern showed evidence of a relationship to when different trials started and ended. Thus, it is apparent that the choice of the window size matters, and also that

different window sizes may provide different levels of information that are equally relevant in understanding the symmetry behavior of the signals.

Symmetry metrics from this study were also informative, albeit sometimes contradictory. Symmetry metrics were broken down by face and house trials, so that differences between these trials could be ascertained. Symmetry proportions suggested that sometimes face trials exhibited more symmetry than house trials, and sometimes they indicated the opposite. Pairwise symmetry proportions also suggested that two regions, DPC and STS, had consistently larger proportions. Symmetry segment lengths for this participant were often found to be longer for house trials than face trials. Model results suggest, however, that no differences existed in the symmetry segment lengths for this participant when $w_{size} = 20$ and $w_{size} = 35$. However, these results should be conducted with a much larger sample size in order to be more informative of these symmetry relationships. Ultimately, these results suggest the process by which symmetry analyses might be conducted to answer questions about functional brain connectivity.

For the drumming study, drumming trios were asked to drum to either a predictable or unpredictable rhythm. Symmetry in their heart rate may indicate how drumming contributes to physiological responses that have been found to correspond to feelings of cohesion. Trivariate windowed cross-correlation plots suggested that meaningful differences may occur in their pattern of association that translate directly to differences in predictable and unpredictable rhythms. Symmetry results revealed more symmetry than expected, but since only one window size was tested, it is possible that a better choice in hyperparameters may produce a more meaningful symmetry pattern, and a more visually-obvious difference in the symmetry of asynchronous and synchronous triads. Symmetry proportions similarly reflect this over-indication of symmetry, and while there were no obvious differences in the proportions of synchronous and asynchronous triads, the highest symmetry proportion was present for a synchronous triad, and the lowest symmetry proportion was present for an asynchronous triad. Symmetry lengths were similarly ambiguous in their interpretation,

including a non-significant difference between synchronous and asynchronous symmetry segment lengths. Testing these analyses on the full sample of drumming triads may provide more insight into these differences. These results demonstrate how the TASS algorithm may be used to answer questions of an interpersonal, social nature.

Limitations

The TASS algorithm has a number of strengths that justify its contribution to the literature on multivariate time series association methods. First, it emphasizes segmentation and the time-based symmetry present between three signals. It can be used to answer questions about the *length* of symmetry. It does so in a way that avoids the assumption of stationarity, and captures non-linear association. Both qualitative and quantitative questions can be tested with the results of the algorithm. It can be applied to data regardless of its content, and therefore has broad applicability. It is modular in its construction and so improvements can be made to each step individually without requiring the whole algorithm be rebuilt.

That being said, there are also a number of limitations that should be discussed. First, the TASS algorithm makes a number of assumptions about the data that may not be ideal for all analyses: 1) that the data be measured on the same time scale; 2) that the data be the same length; 3) that no missingness be present in the data; and 4) that each time series be at least 100 time points long, and ideally much longer than that. For this reason the TASS algorithm may be better suited for physiological data, where missingness is less common, data are generally measured in the same way, and sufficient measurements are often collected. Next, the need for the experimenter to choose hyperparameters requires knowledge of the effect hyperparameters have on symmetry and also knowledge of how quickly time series behavior is expected to change. Additionally, each analysis currently takes roughly 8 hours to run depending on time series length and window size, making analyses with many triads computationally taxing and time consuming.

Various aspects of the TASS algorithm provide limiting factors. The choice of total correlation as the association metric means that time points are indicated as symmetric when only two signals exhibit symmetry, blurring the lines between bivariate and trivariate symmetry. In order to have a clearer picture of what is solely the result of symmetry in all three signals, another information theoretic association metric may be used. Additionally, while the TASS algorithm allows for a finer picture of the symmetry dynamics in three time series, it is limited in its flexibility in two key areas: the window size and the subsequent value. Currently one window size is chosen for the entirety of the time series, which means the experimenter must choose a single length of time for which the given time series are expected to remain the same. It is likely, however, that the perfect window size might vary across the time series. Similarly, the subsequent value used in the analysis is an aggregate of the subsequent values for each row of the CWCC matrix. This is so the prospective and retrospective matrices can be rectangular. This means that the algorithm may over- and under-estimate the number of time points that are relevant for inclusion in the prospective and retrospective matrices, which has the potential to influence how symmetry is indicated, especially for signals with different kinds of symmetry.

Squared Mahalanobis distance (SMD) as the threshold used to distinguish symmetry from non-symmetry is also deserving of close inspection. To begin with, a single threshold is used for all SMDs. This contributes to potential mis-identification of symmetry depending on the scale of the SMD values, and also depending on whether the SMD values provide enough distinction between high and low values where symmetry is and is not present. Part of this problem may be a result of the fact that SMD compares all time points to a single multivariate center, instead of isolating clusters of related time points. Further, the TASS algorithm is currently designed to distinguish symmetric time points from non-symmetric time points. This dichotomous definition of symmetry may not be the most informative for symmetry relationships. It may be better, for instance, to include the possibility for symmetry formation and symmetry breaking as segments.

Lastly, it is apparent that there are some instances where symmetry should be indicated, but is not. This is especially the case for differences in frequency, as seen in Ex9. If there are big enough differences in frequency, especially between periodic signals, symmetry is not indicated. However, there is some debate about whether or not large differences in frequency should still be considered symmetric. On one hand, two sine curves that have different frequencies are still similar, in that the behavior of one corresponds to the behavior of the other. On the other hand, the most important dynamics relevant to the TASS algorithm are local dynamics, and if at short time scales the behavior of one signal is not similar to the behavior of another signal, it should not be indicated as symmetric. A frequency threshold, beyond which two or more signals are not expected to show association, may prove most informative in understanding this distinction. Other instances of incorrect symmetry indication include short symmetry segments, which likely corresponds to the size of the window, and symmetry segments of different lengths, which likely corresponds to the single SMD threshold. It is also worth noting that the results of the applied data analyses were inconclusive due to small sample sizes. While sufficient in demonstrating the utility of the TASS algorithm, it will be necessary to analyze these data with larger sample sizes in order to make well-founded claims about their symmetry patterns.

Future Directions

This work represents the creation and initial testing of the TASS algorithm. More work is needed to understand how well the TASS algorithm captures symmetry, improve upon its functionality in order to increase accuracy, and apply the TASS algorithm to new empirical problems. First, future work should tackle the missingness problem, and allow for missingness to be present in any time series. This will require a system which accounts for the potential bias while also allowing the algorithm to indicate a missing time point as either symmetric or non-symmetric. Speeding up the computation time and allowing for

shorter signals should provide experimenters using the TASS algorithm to conduct analyses faster and with less data needed. Currently the TASS algorithm takes a considerable amount of computation time because of the WCC3 method. The WCC3 method can be improved in a number of ways, including both the way it is constructed and also which program it uses for its computations. Shorter time series might be possible if the algorithm is altered to use more of the surrounding data in its calculation of information.

More work needs to be done in understanding the relationships between the hyperparameters, w_{size} and τ_{max}, and their relationship to symmetry. It appears that larger windows may have more difficulty indicating that time points are symmetric regardless of the data present in each time series. Additionally, it is not always clear whether the window size or the maximum lag should be changed to accommodate for differences in time series. The process of choosing these hyperparameters should be standardized so that researchers have, at the very least, guidelines for how to choose hyperparameter values, or, at the very most, the process of choosing hyperparameter values integrated into the analysis. Conversely, it may be more informative to look at the results conducted with multiple window sizes and maximum lags, rather than selecting a single value for each. In this way it may be possible to look at different scales of association. Further research should investigate how well different hyperparameter values reflect different scales of association. New symmetry metrics may be developed that capture these differences. These new symmetry metrics could provide further nuance to the picture of symmetry that the TASS algorithm captures.

The WCC3 method was developed for this algorithm, and is therefore as of yet unexplored. It may be possible to extract meaningful patterns out of these WCC3 tensors. This study discussed some of the possible interpretations of behaviors in WCC3 matrix slices. Other edifying simulation examples might produce new and informative patterns of association. It is also worth noting that the patterns found in this study were only in WCC3 matrix slices, and not in the tensor as a whole. Further studies might consider

particular patterns present in the WCC3 tensor that incorporate changes in the time domain, across multiple WCC3 matrix slices. Machine learning might be best suited for this task.

The WCC3 method currently uses total correlation as its measure of association. This study covered a number of other potential mutual information metrics that could be used in place of total correlation, that might get at different aspects of the association between windows in three signals. It might be ideal for future studies to use a measure that only captures threeway associations, and not also pairwise relationships like total correlation does. This will make TASS algorithm results solely a reflection of threeway relationships; future work should then compare these results with the sum of all pairwise results to see if the TASS algorithm produces results unique from the pairwise analyses or if the TASS algorithm is simply a sum of the pairwise results. In order to conduct a proper comparison of threeway and pairwise results, it is necessary to develop a bivariate windowed cross-correlation method that uses mutual information as its association metric. A comparison of threeway and pairwise results where both use information-based association metrics would guarantee that any difference in these results was not due to differences in the association metrics.

The TASS algorithm currently uses squared Mahalanobis distance (SMD) as its multivariate distance threshold. As stated previously, SMD is limited in that it compares all time points to a single multivariate center. Future work should consider multivariate distance thresholds that consider the clustering of time points in their similarity, thereby allowing for multiple symmetry segments of different kinds to be uniquely represented in the process of thresholding. Studies that retain SMD as a multivariate distance threshold should consider implementing a systematic approach to choosing a threshold that is specific to the time series at hand. The averaged subsequent value should also be investigated. If an approach to calculating a multivariate distance threshold could be calculated without averaging all subsequent values for each time point, then the average

could be avoided altogether. This would prevent the algorithm from including too many or too few future time points in its calculation for multivariate distance.

Currently the algorithm emphasizes a dichotomous definition of symmetry. The process of thresholding time points into symmetric or non-symmetric may inherently limit a proper understanding of the symmetry behavior. Perhaps the most obvious additions to symmetric and non-symmetric segments would be symmetry formation segments and symmetry breaking segments. Including these as possible segments would give researchers an understanding of when three signals began and end the process of symmetry, and for how long it took for symmetry to be reached and for how long it took for symmetry to be broken. Incorporating these segments would require decisions to be made about the expectations for symmetry segment patterns, and whether or not a non-symmetry segment always separated a symmetry breaking segment and a symmetry formation segment. Future work may consider integrating information about symmetry formation and symmetry breaking behavior.

In this study, the TASS algorithm was applied to two physiological datasets: bold signals from a facial recognition experiment, and heart rates from a drumming experiment. While it was demonstrated how the TASS algorithm could be applied to these empirical studies, only a small subset of the data was analyzed. Future work should conduct full analyses on these data. The TASS algorithm should also be applied to other datasets, particularly those with behavioral data and psychometric data, so that we can understand how the TASS algorithm performs under new kinds of data and new experimental paradigms. The hope is that the TASS algorithm can be used to understand the symmetry segmentation patterns that exist between three time series in meaningful ways.

References

Abdallah, S., & Plumbley, M. (2009). Information dynamics: Patterns of expectation and surprise in the perception of music. *Connection Science*, *21*(2-3), 89–117. doi: 10.1080/09540090902733756

Abney, D. H., Paxton, A., Dale, R., & Kello, C. T. (2014). Complexity matching in dyadic conversation. *Journal of Experimental Psychology: General*, *143*(6), 2304. doi: 10.1037/xge0000021

Anastassiou, D. (2007). Computational analysis of the synergy among multiple interacting genes. *Molecular Systems Biology*, *3*(83). doi: 10.1038/msb4100124

Bakhshayesh, H., Fitzgibbon, S. P., Janani, A. S., Grummett, T. S., & Pope, K. J. (2019). Detecting synchrony in EEG: A comparative study of functional connectivity measures. *Computers in Biology and Medicine*, *105*, 1–15. doi: 10.1016/j.compbiomed.2018.12.005

Barbalet, J. M. (1999). Boredom and social meaning. *British Journal of Sociology*, *50*(4), 631–646. doi: 10.1111/j.1468-4446.1999.00631.x

Barlow, H. B., & Reeves, B. C. (1979). The versatility and absolute efficiency of detecting mirror symmetry in random dot displays. *Vision Research*, *19*(7), 783–793. doi: 10.1016/0042-6989(79)90154-8

Ben-Gal, I. (2010). Outlier Detection. In O. Maimon & L. Rockach (Eds.), *Data mining and knowledge discovery handbook: A complete guide for practitioners and researchers* (pp. 131–142). New York: Springer.

Bertamini, M., Friedenberg, J. D., & Kubovy, M. (1997). Detection of symmetry and perceptual organization: The way a lock-and-key process works. *Acta Psychologica*, *95*(2), 119–140. doi: 10.1016/s0001-6918(96)00038-8

Blairy, S., Herrera, P., & Hess, U. (1999). Mimicry and the judgment of emotional facial expressions. *Journal of Nonverbal Behavior*, *23*(1), 5–41. doi: 10.1023/A

Boker, S. M., Cohn, J. F., Theobald, B.-J., Matthews, I., Brick, T. R., & Spies, J. R.

(2009). Effects of damping head movement and facial expression in dyadic conversation using real-time facial expression tracking and synthesized avatars. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1535), 3485–3495. doi: 10.1098/rstb.2009.0152

Boker, S. M., Cohn, J. F., Theobald, B.-J., Matthews, I., Mangini, M., Spies, J. R., ... Brick, T. R. (2011). Something in the way we move: Motion dynamics, not perceived sex, influence head movements in conversation. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(3), 874–91. doi: 10.1037/a0021928

Boker, S. M., Deboeck, P. R., Edler, C., & Keel, P. K. (2010). Generalized local linear approximation of derivatives from time series. In S.-M. Chow & E. Ferrar (Eds.), *Statistical methods for modeling human dynamics: An interdisciplinary dialogue* (pp. 179–212). Boca Raton, FL: Taylor & Francis. doi: 10.4324/9780203864746

Boker, S. M., & Kubovy, M. (1998). Predicting the grouping of rhythmic sequences using local estimators of information content. In D. Rosenthal & H. G. Okuno (Eds.), *Computational auditory scene analysis* (pp. 309–320). Hillsdale NJ: Lawrence Ehrlbaum.

Boker, S. M., & Rotondo, J. L. (2002). Symmetry building and symmetry breaking in synchronized movement. In M. Stamenov & V. Gallese (Eds.), *Mirror neurons and the evolution of brain and language.* (pp. 163–171). Philadelphia: John Benjamins Publishing. doi: 10.1016/S0375-9474(01)00406-7

Boker, S. M., Rotondo, J. L., Xu, M., & King, K. (2002). Windowed cross-correlation and peak picking for the analysis of variability in the association between behavioral time series. *Psychological Methods*, *7*(3), 338–355. doi: 10.1037/1082-989X.7.3.338

Boly, M., Perlbarg, V., Marrelec, G., Schabus, M., Laureys, S., Doyon, J., ... Benali, H. (2012). Hierarchical clustering of brain activity during human nonrapid eye movement sleep. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(15), 5856–5861. doi: 10.1073/pnas.1111133109

Bornstein, M. H., Ferdinandsen, K., & Gross, C. G. (1981). Perception of symmetry in infancy. *Developmental Psychology*, *17*(1), 82–86. doi: 10.1037/0012-1649.17.1.82

Butler, E. A. (2011). Temporal interpersonal emotion systems: The "TIES" that form relationships. *Personality and Social Psychology Review*, *15*(4), 367–393. doi: 10.1177/1088868311411164

Butler, E. A., & Randall, A. K. (2013). Emotional coregulation in close relationships. *Emotion Review*, *5*(2), 202–210. doi: 10.1177/1754073912451630

Cappella, J. N. (1997). Behavioral and judged coordination in adult informal social interactions: Vocal and kinesic indicators. *Journal of Personality and Social Psychology*, *72*(1), 119–131. doi: 10.1037/0022-3514.72.1.119

Chartrand, T. L., & Bargh, J. A. (1999). The Chameleon Effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology*, *76*(6), 893–910.

Choukroune, P., Gapais, D., & Merle, O. (1987). Shear criteria and structural symmetry. *Journal of Structural Geology*, *9*(5-6), 525–530. doi: 10.1016/0191-8141(87)90137-4

Codrons, E., Bernardi, N. F., Vandoni, M., & Bernardi, L. (2014). Spontaneous group synchronization of movements and respiratory rhythms. *PLoS ONE*, *9*(9). doi: 10.1371/journal.pone.0107538

Dauwels, J., Vialatte, F., Musha, T., & Cichocki, A. (2010). A comparative study of synchrony measures for the early diagnosis of Alzheimer's disease based on EEG. *NeuroImage*, *49*(1), 668–693. doi: 10.1016/j.neuroimage.2009.06.056

Fitzpatrick, J., Gareau, A., Lafontaine, M.-F., & Gaudreau, P. (2016). How to use the Actor-Partner Interdependence Model (APIM) to estimate different dyadic patterns in MPLUS: A step-by-step tutorial. *The Quantitative Methods for Psychology*, *12*(1), 74–86. doi: 10.20982/tqmp.12.1.p074

Garner, W. (1962). *Uncertainty and Structure as Psychological Concepts*. New York: John Wiley & Sons, Inc.

Gordon, I., Gilboa, A., Cohen, S., Milstein, N., Haimovich, N., Pinhasi, S., & Siegman, S. (2020). Physiological and Behavioral Synchrony Predict Group Cohesion and Performance. *Scientific Reports*, *10*(8484), 1–12. doi: 10.1038/s41598-020-65670-1

Gross, D. J. (1996). The role of symmetry in fundamental physics. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(25), 14256–14259. doi: 10.1073/pnas.93.25.14256

Hanel, P. H., Maio, G. R., & Manstead, A. S. (2019). A new way to look at the data: Similarities between groups of people are large and important. *Journal of Personality and Social Psychology*, *116*(4), 541–562. doi: 10.1037/pspi0000154

Hendry, D. F., & Juselius, K. (2000). Explaining cointegration analysis: Part 1. *The Energy Journal*, *21*(1), 1–42.

Honey, C. J., Kötter, R., Breakspear, M., & Sporns, O. (2007). Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(24), 10240–10245. doi: 10.1073/pnas.0701519104

Ivanov, P. C., Rosenblum, M. G., Peng, C. K., Mietus, J., Havlin, S., Stanley, H. E., & Goldberger, A. L. (1996). Scaling behaviour of heartbeat intervals obtained by wavelet-based time-series analysis. *Nature*, *383*(26), 323–327. doi: 10.1038/383323a0

Jaynes, E. T. (1963). Information Theory and Statistical Mechanics. In K. W. Ford (Ed.), *Brandeis lectures in theoretical physics, volume 3: Statistical physics* (pp. 181–219). New York: W. A. Benjamin, Inc.

Jaynes, E. T. (1968). Prior Probabilities. *IEEE Transactions on Systems Science and Cybernetics*, *4*(3), 227–241. doi: 10.1109/TSSC.1968.300117

Keller, P. E., Novembre, G., & Hove, M. J. (2014). Rhythm in joint action: Psychological and neurophysiological mechanisms for real-time interpersonal coordination. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *369*. doi: 10.1098/rstb.2013.0394

Keogh, E., Chu, S., Hart, D., & Pazzani, M. (2004). Segmenting time series: A survey and novel approach. In M. Last, A. Kandel, & H. Bunke (Eds.), *Data mining in time series databases* (pp. 1–23). Singapore: World Scientific Publishing Co.

Koban, L., Ramamoorthy, A., & Konvalinka, I. (2019). Why do we fall into sync with others? Interpersonal synchronization and the brain's optimization principle. *Social Neuroscience*, *14*(1), 1–9. doi: 10.1080/17470919.2017.1400463

Le Page, Y. (1987). Computer derivation of the symmetry elements implied in a structure description. *Journal of Applied Crystallography*, *20*(3), 264–269. doi: 10.1107/S0021889887086710

Leys, C., Delacre, M., Mora, Y. L., Lakens, D., & Ley, C. (2019). How to classify, detect, and manage univariate and multivariate outliers, with emphasis on pre-registration. *International Review of Social Psychology*, *32*(1), 1–10. doi: 10.5334/irsp.289

Magyari, L., Bastiaansen, M. C. M., de Ruiter, J. P., & Levinson, S. C. (2014). Early anticipation lies behind the speed of response in conversation. *Journal of Cognitive Neuroscience*, *26*(11), 2530–2539.

Mahalanobis, P. C. (1936). On the generalized distance in statistics. *Proceedings of the National Institute of Sciences of India*, 49–55.

Michel, L. (1980). Symmetry defects and broken symmetry. *Reviews of Modern Physics*, *52*(3), 617–651.

Mikosch, T., & Starica, C. (2004). Nonstationarities in financial time series, the long-range dependence, and the IGARCH effects. *Review of Economics and Statistics*, *86*(1), 378–390. doi: 10.1162/003465304323023886

Moeck, P. (2019). On classification approaches for crystallographic symmetries of noisy 2D periodic patterns. *IEEE Transactions on Nanotechnology*, *18*, 1166–1173.

Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement*, *2*(4), 201–218.

Moulder, R. G., Boker, S. M., Ramseyer, F., & Tschacher, W. (2018). Determining synchrony between behavioral time series: An application of surrogate data generation for establishing falsifiable null-hypotheses. *Psychological Methods*, *23*(4), 757–773. doi: 10.1037/met0000172

Nelson, C. R., & Plosser, C. I. (1982). Trends and random walks in macroeconomic time series: Some evidence and implications. *Journal of Monetary Economics*, *10*, 139–162. doi: 10.1016/j.jmacro.2011.10.001

Oglesby, D. D., Archuleta, R. J., & Nielsen, S. B. (1998). Earthquakes on dipping faults: The effects of broken symmetry. *Science*, *280*(5366), 1055–1059. doi: 10.1126/science.280.5366.1055

Oullier, O., de Guzman, G., Jantzen, K. J., Lagarde, J., & Scott Kelso, J. A. (2008). Social coordination dynamics: Measuring human bonding. *Social Neuroscience*, *3*(2), 178–192. doi: 10.1080/17470910701563392

Pais, A. (1966). Dynamical symmetry in particle physics. *Reviews of Modern Physics*, *38*(2), 215–255. doi: 10.1103/RevModPhys.38.215

Pillai, A. S., & Jirsa, V. K. (2017). Symmetry breaking in space-time hierarchies shapes brain dynamics and behavior. *Neuron*, *94*(5), 1010–1026. doi: 10.1016/j.neuron.2017.05.013

Rennung, M., & Göritz, A. S. (2016). Prosocial consequences of interpersonal synchrony: A Meta-Analysis. *Zeitschrift fur Psychologie / Journal of Psychology*, *224*(3), 168–189. doi: 10.1027/2151-2604/a000252

Reutzel-Edens, S. M. (2020). Symmetry in the making. *Nature Chemistry*, *12*(10), 883–890. doi: 10.1038/s41557-020-0523-3

Rosen, J. (1983). *A Symmetry Primer for Scientists*. New York: John Wiley & Sons, Inc.

Saspiturry, N., Razin, P., Baudin, T., Serrano, O., Issautier, B., Lasseur, E., ... Leleu, S. (2019). Symmetry vs. asymmetry of a hyper-thinned rift: Example of the Mauléon Basin (Western Pyrenees, France). *Marine and Petroleum Geology*, *104*, 86–105. doi:

10.1016/j.marpetgeo.2019.03.031

Schmidt, R. C., & Richardson, M. J. (2008). Dynamics of interpersonal coordination. In *Coordination: Neural, behavioral, and social dynamics* (pp. 281–308). Berlin: Springer.

Schneidman, E., Still, S., Berry, M. J., & Bialek, W. (2003). Network information and connected correlations. *Physical Review Letters*, *91*(23), 3–6. doi: 10.1103/PhysRevLett.91.238701

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*(4), 379–423, 623–656. doi: 10.1002/j.1538-7305.1948.tb00917.x

Simon, H. A. (1952). A formal theory of interaction in social groups. *American Sociological Review*, *17*(2), 202–211.

Sjobeck, G. R., Boker, S. M., Scheidt, C. E., & Tschacher, W. (n.d.). The Pairwise Approximate Spatiotemporal Symmetry (PASS) Algorithm: A method for segmenting time series pairs. *Manuscript in preparation.*

Smith, S. W. (1997). *The Scientist and Engineer's Guide to Digital Signal Processing.* San Diego: California Technical Publishing.

Spek, A. L. (2003). Single-crystal structure validation with the program PLATON. *Journal of Applied Crystallography*, *36*(1), 7–13. doi: 10.1107/S0021889802022112

Strogatz, S. (2003). *Sync: How Order Emerges from Chaos in the Universe, Nature, and Daily Life.* New York: Hatchette Books.

Turchin, P., & Taylor, A. D. (1992). Complex dynamics in ecological time series. *Ecology*, *73*(1), 289–305.

Wagner, N., & Ashkenasy, G. (2009). Symmetry and order in systems chemistry. *Journal of Chemical Physics*, *130*(164907), 1–6. doi: 10.1063/1.3118649

Watanabe, S. (1960). Information theoretical analysis of multivariate Correlation.pdf. *IBM Journal of Research and Development*, *4*, 66–82.

Zabrodsky, H., Peleg, S., & Avnir, D. (1992). Continuous symmetry measures. *Journal of*

the American Chemical Society, *114*(20), 7843–7851. doi: 10.1021/ja00046a033