Using Person-Specific Neural Networks to Characterize Heterogeneity in Eating Disorders:

Illustrative Links Between Emotional Eating and Ovarian Hormones

Adriene M. Beltz

University of Michigan

Jason S. Moser, David C. Zhu, S. Alexandra Burt, and Kelly L. Klump

Michigan State University

Corresponding author:

Adriene M. Beltz

Department of Psychology

2227 East Hall

530 Church Street

Ann Arbor, MI 48109

734-647-6788

abeltz@umich.edu

Abstract

Objective: Emotional eating has been linked to ovarian hormone functioning, but no studies to-date have considered the role of brain function. This knowledge gap may stem from methodological challenges: Data are heterogeneous, violating assumptions of homogeneity made by between-subjects analyses. The primary aim of this paper is to describe an innovative within-subjects analysis that models heterogeneity and has potential for filling knowledge gaps in eating disorder research. We illustrate its utility in an application to pilot neuroimaging, hormone, and emotional eating data across the menstrual cycle. **Methods**: Group iterative multiple model estimation (GIMME) is a person-specific network approach for estimating sample-, subgroup-, and individual-level connections between brain regions. To illustrate its potential for eating disorder research, we apply it to pilot data from 10 female twins (N=5 pairs) discordant for emotional eating and/or anxiety, who provided two resting state fMRI scans and hormone assays. We then demonstrate how the multimodal data can be linked in multilevel models. **Results**: GIMME generated person-specific neural networks that contained connections common across the sample, shared between co-twins, and unique to individuals. Illustrative analyses revealed positive relations between hormones and default mode connectivity strength for control twins, but no relations for their co-twins who engage in emotional eating or who had anxiety. **Discussion**: This paper showcases the value of person-specific neuroimaging network analyses and its multimodal associations in the study of heterogeneous biopsychosocial phenomena, such as eating behavior.

keywords: connectivity, emotional eating, estrogen, heterogeneity, person-specific, precision healthcare, progesterone, resting state, twin study

No two individuals diagnosed with mental health conditions are the same. Individuals vary in disease etiology, presentation, prognosis, and treatment effectiveness, owing to their unique histories, comorbidities, and biological makeups (Sysko, Hildebrandt, Wilson, Wilfley, & Agras, 2010). Yet, heterogeneous individuals are often assumed to be homogeneous in research (e.g., through pooling) and clinic (e.g., in treatment) settings. This mismatch between an individual's experience of a disorder and the characterization of it by professionals on the bench and at the bedside likely contributes to suboptimal etiologic theories and long-term outcomes. The aim of this paper is to address this mismatch by introducing and illustratively applying to variable emotional eating data an innovative analysis technique - group iterative multiple model estimation (GIMME; Beltz & Molenaar, 2016; Gates & Molenaar, 2012) - that capitalizes on heterogeneity to identify person-specific results.

Between-subjects analyses dominate clinical research. These traditional statistical approaches (e.g., regression) average across *inter-individual variation* to generate inferences that apply to each individual in a sample and to similar others who were not directly sampled (Cattell, 1952). They assume that variability between people is random and will cancel out. This approach has incredible utility for describing a population and has led to monumental basic and applied clinical insights. But, it comes with two under-appreciated, related caveats. First, samples must be homogeneous, otherwise, average results may not apply to anyone in the sample; this assumption is specified in the ergodic theorem (Birkhoff, 1931; Molenaar, 2004). Second, "average" people do not walk into clinics; individuals do. Treatments developed from homogenous sample averages may contribute to poor responses from heterogeneous individuals, highlighting the need for precision healthcare (Davidson & Cheung, 2017; Gambhir, Ge, Vermesh, & Spitler, 2018).

An alternative to between-subjects analyses are within-subjects analyses (Birkhoff, 1931; Cattell, 1952; Molenaar, 2004). These statistical approaches reveal patterns in *intra-individual variation* reflected in intensive longitudinal or time series data and permit inferences that apply uniquely to an individual. They assume that variability between people is meaningful. Within-subjects analyses complement between-subjects analyses by enabling researchers to ask and answer different kinds of research questions about the *person-specific* interplay among biopsychosocial variables that contribute to a disorder's etiology, presentation, and treatment. They consider each person as a sample of size one.

Focusing on the individual ensures accurate statistical results when studying phenomena that vary across people and time, including many eating disorders. The eating disorder group is often more heterogeneous than the control group in clinical research (e.g., in neural signals; Bohon & Stice, 2011; Marsh et al., 2009), and a focus on mean-level differences might produce biased estimates because the average is not representative (Molenaar, 2004). Focusing on the individual is also consistent with the rising prominence of precision healthcare (Davidson & Cheung, 2017; Gambhir et al., 2018), and has demonstrated importance in clinically-relevant phenomena. In personality research, person-specific analyses have shown that the common five-factor model explaining variation between people only applies to 14% of individuals (Molenaar & Campbell, 2009).

Despite their value, there is a paucity of sophisticated analytic tools for conducting within-subjects analyses. One promising technique that even affords group-level inferences, so that focus on a person does not prevent generalizations across people, is GIMME (Gates & Molenaar, 2012). It is a network analysis approach that identifies statistically prominent relations among variables in a system in order to explain time-indexed patterns in the observed "signals"

of those variables. In fMRI research, the variables are brain regions of interest (ROIs) thought to work as a coordinated network, and blood oxygen-level dependent (BOLD) signal is used to measure neuronal activity. GIMME is a person-specific approach, so it is not challenged by, but rather leverages, heterogeneity; it does not average or pass variance between levels of analysis even though it identifies network features common across people (see Methods).

Importantly, unlike controversial cross-sectional psychopathology networks (Borsboom & Cramer, 2013; Forbes, Wright, Markon, & Krueger, 2017), GIMME uses time series data to create *temporal networks* that include both individual-level (person-specific) and group-level (sample-relevant) information. Functional magnetic resonance imaging (fMRI) data are a prime example. In this domain, GIMME's accuracy and precision have been demonstrated (Gates & Molenaar, 2012). It performed well in large-scale simulations in which temporal patterns of connections between different numbers of brain regions with varying characteristics (e.g., noise) were generated in order to determine how well different network analysis techniques could recover the "true" patterns. GIMME recovered more true connections and identified fewer false positives than most other approaches, including (partial) correlations, Granger causality, and Bayesian nets (Gates & Molenaar, 2012; Smith et al., 2011). Some of these approaches are common in eating disorder research, and their features in comparison to GIMME are summarized in Table 1. As shown, GIMME is unique in modeling contemporaneous and lagged connections as well as individual- and group-level connections.

Despite its promise, GIMME has not yet been used in eating disorder research, but it has provided insight in studies of other phenotypes (e.g., addiction and depression; Beltz et al., 2013; Price et al., 2017). Thus, the goal of the current manuscript is to introduce GIMME to the eating disorder community as a tool for person-specific analysis of fMRI data. Its premise and

mathematics will be reviewed, and then its value for multimodal eating disorder research will be illustrated in a heterogeneous, pilot data set consisting of resting state brain function assessed at two timepoints across the menstrual cycle in female twin pairs discordant for emotional eating and/or anxiety. These pilot data were ideally suited to GIMME: Clinical status (i.e., high versus low on emotional eating and/or anxiety) was assessed prior to intake, ovarian hormones and emotional eating were assessed on the same day two fMRI scans were conducted, and the use of discordant twin pairs provided both cases and controls.

Following person-specific GIMME analyses, the statistical model implemented by GIMME was also applied to the sample average-level fMRI data. In other words, a network analysis of between-subjects, inter-individual variation was conducted to compare to GIMME's within-subjects analysis of intra-individual variation. Since pilot data are heterogeneous, differences between average-level and GIMME results showcase how inaccurate findings can result from inappropriately averaging across people and time.

Methods

GIMME methods are described in detail first, followed by a brief description of neuroimaging and behavioral data methods. See Supplemental Material for extensive descriptions.

GIMME

Conceptual Premise. GIMME boasts several characteristics for the analysis of heterogeneous data. First, it is a person-specific approach that incorporates group-level information. Results provide a network for each individual, but some network connections are common to everyone in the sample, and some are unique to the individual. Consider, for example, a sample of N = 5 and a network with four ROIs: A, B, C, and D. GIMME would

indicate that all 5 people have a connection between A and B, but that only one person has a connection between A and C.

Second, GIMME implements unified structural equation models (uSEMs; Gates, Molenaar, Hillary, Ram, & Rovine, 2010), so connections are directed, reflecting statistical prediction from one ROI to another (or to itself), with a positive or negative magnitude. In the example, GIMME would indicate that the person-specific connection between A and C goes *from* A *to* C and is positive with a standardized β of .43.

Third, connections reflect temporal information. Some connections are contemporaneous, wherein signal in one ROI predicts signal in another ROI at the same time point (i.e., measurement occasion, such as an fMRI volume). Other connections are lagged, wherein signal in one ROI predicts signal in another ROI (or itself) at the next time point. In the example, the directed connection between A and C is lagged, such that signal in A predicts signal in C at the next volume.

Fourth, data-driven versions of GIMME have been fully automated in R and Matlab® (Beltz & Molenaar, 2016; Lane, Gates, & Molenaar, 2017). They produce sparse (not saturated) networks that contain only the connections required to fit the observed variation among all ROIs. GIMME begins with a null network and iteratively adds connections that account for the most signal. Once enough signal is accounted for (according to model fit indices), GIMME stops. In the example, GIMME would not estimate all possible pairs of relations among A, B, C, and D, but only those required to explain the temporal variation among them for each person.

Application to Resting State Brain Function. Resting state brain function (i.e., neural activity when not engaged in a task) reflects the brain's physiological and psychological baseline (Gusnard & Raichle, 2001) and serves as a marker for neuropsychiatric disease (Fox & Greicius,

2010). It is typically understood in terms of synchronized patterns of connectivity that reflect networks among integrated brain regions - not localized activity, which predominates task fMRI.

An ROI time series (i.e., a matrix of 10 ROIs as the columns by 280 volumes as the rows for each participant - see Illustrative Data and Supplemental Material) were submitted to a Matlab® version of GIMME (Beltz & Molenaar, 2016) modified to estimate twin pair-level (within each family) as well as group-level (everyone in the sample) and individual-level (one scan for one twin) connections. This is defined as:

$$\boldsymbol{\eta}_i(t) = (\mathbf{A}_i + \mathbf{A}_i^f + \mathbf{A}_i^g)\boldsymbol{\eta}_i(t) + (\boldsymbol{\Phi}_{1,i} + \boldsymbol{\Phi}_{1,i}^f + \boldsymbol{\Phi}_{1,i}^g)\boldsymbol{\eta}_i(t-1) + \boldsymbol{\zeta}_i(t),$$

where $\boldsymbol{\eta}$ is the observed 10-ROI time series for individual $i = 1, 2, 3, \ldots, 20$ at time $t = 1, 2, 3, \ldots 280$. $\mathbf{A}$ is the 10 x 10 matrix of directed contemporaneous ROI connections, with a diagonal fixed to 0; it is a structural equation model within GIMME. $\boldsymbol{\Phi}$ is the 10 x 10 matrix of directed lagged ROI connections, with estimable autoregressive connections on the diagonal (ROIs can predict themselves at the next time point); it is a vector autoregressive model within GIMME. $\boldsymbol{\zeta}$ is the error, with a zero mean, diagonal covariance matrix, and no sequential dependencies. Superscripts reflect connections that are estimated for individuals in the same family $f = 1, 2, 3, 4, 5$ or for everyone in the group g (sample). The subscript i, reflecting connections estimated for an individual at a given scan, co-occurs with f and g because all connections have person-specific estimates (β's). Details are in a recent tutorial (Beltz & Gates, 2017).

GIMME generates person-specific networks in several steps. First, a null model (without contemporaneous or lagged connections) is fit to each person's data. Sometimes the null model contains the autoregressive connections in $\boldsymbol{\Phi}$ (used here) because they reflect stability in time series with relatively short measurement intervals. Second, a group-level model is identified. Lagrange Multiplier equivalents (modification indices; Sörbom, 1989) are used to determine

which connection in **A** or **Φ**, if freed, would significantly and maximally improve model fit for a criterion (usually 75% of the sample, but 100% was used here due to the small sample). That connection is added to everybody's model, and models are re-estimated. This process iterates until no more connections meet the criterion. Third, family-level models are identified for each twin pair. The group-level model is considered the "null," and modification indices are iteratively used to determine which connection, if freed, would significantly and maximally improve model fit for the networks of both twins. Fourth, individual-level models are identified. The combined group- and family-level model is considered the "null," and modification indices determine which connection, if freed, would significantly and maximally improve model fit for an individual, iterating until the model fits well according to at least two of four indices (Brown, 2006): comparative fit index (CFI) $\geq$ .95; non-normed fit index (NNFI) $\geq$ .95; root mean squared error of approximation (RMSEA) $\leq$ .05; standardized root mean residual (SRMR) $\leq$ .05. Thus, GIMME results provide a (potentially) unique model for each person in the sample. Some connections are estimated for everybody, some for certain twin pairs, and some only for individual people or scans, and models are evaluated with standard fit indices. All connections have a β (with significance test) estimated at the individual-level; in other words, the group- and family-level parameters do not reflect averages across the sample or twin pairs, respectively, but rather the structure of fixed and freely estimable connections.

Finally, an average-level uSEM analysis (Gates et al., 2010) was conducted on the mean 10 (ROIs) x 280 (volumes) matrix across participants and scans. This created a between-subjects network based on the assumption of homogeneity. The data-driven model was fit in LISREL (a dependency of the Matlab® version of GIMME) using a null model that contained

autoregressive parameters. Connections were iteratively added (using modification indices) until the model fit well according to two of four indices listed above.

Illustrative Data (see Supplemental Material for full methods)

Pilot data were selected to capitalize on the heterogeneity inherent in clinical status and ovarian hormone levels to showcase GIMME and its potential for the eating disorder community. Participants included 10 female twins from 5 twin pairs (6 monozygotic, 4 dizygotic) ages 18-29 (M = 22.22, SD = 4.46) recruited from the *Twin Study of Hormones and Behavior across the Menstrual Cycle* (TSHBMC) from the Michigan State University Twin Registry. Because the original aims of this pilot study were to examine transdiagnostic factors contributing to binge eating and anxiety, the sample was preferentially selected to be discordant on emotional eating and/or anxiety levels; see Supplementary Material for definitions of discordance. Three pairs (60%) were discordant for emotional eating, one pair (20%) for emotional eating and anxiety, and one pair (20%) for anxiety, but this pair also showed a marked difference in emotional eating across the study (case: M = .31, SD = .17; control: M = .00, SD = .02), so was included in analyses.

Participants collected salivary samples of estrogen and progesterone and provided emotional eating data (via the Dutch Eating Behavior Questionnaire; Van Strien, Frijters, Bergers, & Defares, 1986) daily for 35 consecutive days using previously described protocols (Klump et al., 2013). Past studies in humans show significant associations between emotional eating and higher levels of estradiol and progesterone (for a review, see Klump, Culbert, & Sisk, 2017), and animal studies provide strong evidence in favor of ovarian hormone regulation of food intake (Asarian & Geary, 2013) and reward-related neural systems (Becker, 2009). Salivary hormone samples were assayed using standard enzyme immunoassay kits, and hormone levels

were quantified using five-day rolling averages (Klump et al., 2013). Hormone levels <u>on scan days only</u> were used in analyses.

We used previously-established procedures (Klump et al., 2015) and ovulation kits to ensure different hormone levels at each of two scans conducted during the 35-day study. Procedures successfully captured pre- versus post-ovulation for six (60%) participants. Two (20%) participants had scans during anovulatory cycles, and two (20%) were in pre-ovulation at both scans. Hormone profiles of these participants still differed between scans (estrogen difference range: .28-.89; progesterone difference range: 15.83-23.36). All participants were included in analyses because GIMME models heterogeneity, and multilevel models (see Results) accounted for dependencies between scans.

Resting state brain function was assessed with 12-minute fMRI scans using a standard echo planar imaging sequence conducted on a 3 Tesla GE scanner; structural data were also acquired. Standard preprocessing, with motion and noise correction, was conducted, and functional time series from 10 ROIs was extracted for subsequent network analyses. ROIs are listed in Table S1 and displayed in Figure 1. See Supplemental Material for details.

<div align="center">

Results

</div>

GIMME: Application to Resting State Brain Function

GIMME models reflecting within-person variation fit each participant's data well, according to indices reported in Table 2. Figure 2 illustrates GIMME models for one twin pair discordant for emotional eating at pre- and post-ovulation, and Figure 3 shows models for the twin pair discordant for anxiety. All models contain 15 group-level connections (thick black lines): Ten are autoregressions (dashed circular arrows estimated in the null model), and five are contemporaneous, primarily between contralateral (opposite-hemisphere) brain regions; there

was also a contemporaneous connection from the left lateral parietal lobe to the PCC (see ROI acronyms in Figure 1). The group-level connections represent homogeneity across the sample because they were important for all participants' models. Note, however, that the β weights of these connections differ across participants and scans because they are estimated uniquely in each model. This is seen in Table 3 and Figures 2 and 3. For the twin who engages in emotional eating (Figure 2), the connection from the left to the right OFC decreased from .95 to .68 from pre- to post-ovulation, while the same connection increased from .53 to .62 for her control co-twin. For the twin with high anxiety (Figure 3), that same connection increased from .33 to .53 from pre- to post-ovulation, while it remained relatively constant for her control co-twin at .70 and .68. Thus, GIMME determines the group-level structure of the data by fitting the same connections, but calculates person-specific estimates – and not average parameters – of those connections.

As reported in Table 3, GIMME models also contained between 8 and 11 twin pair-level connections (thick gray lines in Figures 2 and 3); the majority were contemporaneous. These connections represent homogeneity *within* genes and families as well as heterogeneity *between* genes and families. They were important for model fit in both scans for particular twin pairs. For instance, in Figure 2, there was one lagged and seven contemporaneous pair-level connections, and they were mainly between ipsilateral (same-hemisphere) ROIs or between medial and lateral ROIs. In Figure 3, there were three lagged and eight contemporaneous pair-level connections, and again, they were mainly between ipsilateral or between medial and lateral ROIs; there were also two feedback loops (i.e., lagged connections between contralateral ROIs that were contemporaneously-connected). Even though pair-level connections seemed to serve a similar

function across pairs (reflecting ipsilateral connectivity), they did so in pair-specific ways, as only four connections were common between participants in Figures 2 and 3.

Finally, all but one GIMME model contained person-specific connections (thin black lines in Figures 2 and 3); there were between 0 and 20 per model ($M = 6.10$, $SD = 5.52$); see Table 3. One scan for one twin did not require any unique connections because the group- and pair-level connections already fit the data well. These connections reflect heterogeneity across individuals because they are needed to explain the data even after homogeneity across all people and co-twins has been modeled. In Figure 2, the twin who engaged in emotional eating had eight connections during pre-ovulation with most being lagged (and two of which were present during post-ovulation), while her co-twin had four connections during pre-ovulation with most being contemporaneous. In Figure 3, the twin with high anxiety had only one person-specific connection during both pre- and post-ovulation; it was the same lagged link between the PCC and RLP. It reflects stability within this participant.

To provide a comparison between person-specific GIMME results and results from a comparable between-subjects analysis that ignores heterogeneity, a uSEM was fit to sample average-level data. The network is shown in Figure 4, and it fit the data well: $\chi^2(124) = 295.11$, $p < .001$, CFI = .97, NNFI = .95, RMSEA = .07, SRMR = . 06. Consistent with the group-level GIMME results, most connections were between contralateral ROIs or among ROIs in the established default mode network. There were, however, two contemporaneous connections from the MPFC to left reward network regions (i.e., LS and LOFC) that highlight the shortcomings of the average-level approach. The MPFC-to-LS connection was displayed by only three pairs (including the pair in Figure 2), and the MPFC-to-LOFC connection was displayed two pairs (including the pair in Figure 3). Accuracy was obscured by averaging! The average-level model

assumed that the MPFC was connected moderately to the LS and LOFC in all participants, but in fact, those connections were driven by specific twin pairs and were not applicable to others.

Illustrative Example: Linking Resting State Brain Function to Ovarian Hormones and Emotional Eating

Resting state brain function results from GIMME were linked to ovarian hormones and clinical status (i.e. twin with versus without emotional eating/anxiety) to illustrate how GIMME can be used to identify novel brain-behavior relations. Descriptive statistics for emotional eating, estrogen, and progesterone during pre- and post-ovulation are shown in Table S2. As expected, cases had higher emotional eating scores than controls at both pre- and especially post-ovulation, and both groups showed variability in hormone levels between scans.

Moderation analyses implemented in multilevel models with random intercepts and nesting by scans and twin pair were used to examine between-subject hormone-connectivity relations in cases versus control. Due to their illustrative nature, analyses were not corrected for multiple comparisons. Clinical status, estrogen and progesterone (in separate models), and their interaction were used to predict GIMME model parameters. Many different parameters could be used, but because follow-up analyses are between-subject, we focused on β weights of GIMME-identified group-level connections listed in Table 3; they are uniquely estimated for each participant and scan and are not zero-inflated (as are pair- and individual-level connections).

Results revealed a main effect of clinical status on the LLP-to-PCC connection for estrogen, $b = -.20$, $p = .037$, and progesterone, $b = -.19$, $p = .029$, such that controls had greater connectivity strength than cases. There were also interactions for the RLP-to-LLP connection for estrogen, $b = -.10$, $p = .078$, and progesterone, $b = -.15$, $p = .008$. Strength was positively related to estrogen ($b = .11$) and progesterone ($b = .11$), in controls, but there was little evidence for a

relation in cases ($b = -.01$ and $b = -.03$, respectively). Although results are merely illustrative because of the small sample, some provocative findings emerged: Compared to controls, women who engage in emotional eating or had anxiety showed weaker default mode connectivity (LLP → PCC) and weaker relations between ovarian hormones and default mode connections (RLP → LLP).

Discussion

Our goal was to introduce to the eating disorder community a novel network analysis technique called GIMME that is well-suited to examinations of heterogeneity, and to demonstrate how it can be combined with highly variable clinical and biological data to investigate new research questions with relevance to precision healthcare. Analysis of intentionally heterogeneous illustrative data showed that GIMME can be used to create person-specific models of brain networks that can then be linked to ovarian hormone levels (spanning pre- and post-ovulation) and clinical status (i.e., emotional eating, anxiety, or both). Analysis of these data – after averaging them – using the statistical model implemented by GIMME (mimicking traditional between-subjects analyses of inter-individual variation based on the faulty assumption of homogeneity) produced some spurious findings.

GIMME results (reflecting a within-subjects analyses of intra-individual variation based on the assumption of heterogeneity) revealed several group-level connections, primarily between contralateral ROIs, many of which were detected in the average-level analyses. Thus, there is some evidence for convergence of GIMME and average-level results. Due to functional similarities between contralateral regions, these connections are logical reflections of brain function and expected to replicate across methods and samples.

In order to account for the relatedness of co-twins, GIMME was applied for the first time with an *a priori* family level. Results showed a similar number of twin pair-level connections (between 8 and 11) across the two scans. These connections were *misrepresented by the average-level results*, which indicate that all individuals had positive, moderate associations from the MPFC to the LS and the LOFC. Twin pair-level analyses, however, revealed that these connections were present for some pairs, but not others. Thus, the assumption of homogeneity made in the between-subjects analysis in order to permit generalization across people was violated; results only generalized to select individuals. It is unlikely that power is the primary reason for the inaccurate average-level results because – even with a larger sample size – the average would still collapse across heterogeneous individuals, who vary from each other and across time in emotional eating and anxiety. It is this inappropriate averaging over heterogeneity that leads to inaccuracies (Birkhoff, 1931; Molenaar, 2004).

Finally, GIMME results revealed unique connections for almost all individuals and scans, reflecting heterogeneity across people and time; these connections were absent from the sample-average analysis. Among one twin pair (Figure 2), the twin who engaged in emotional eating showed the most individual-level connections during pre-ovulation, perhaps reflecting increased neural regulation important for reducing risk, but when that cross-talk ends (e.g., due to hormone modulation during post-ovulation), emotional eating ensues. Thus, results suggest a potential person-specific clinical target of increasing regulation through neural connectivity (e.g., via neurofeedback) during risk episodes. A distinct example is provided by another twin pair (Figure 3) in which the twin discordant for anxiety did not show remarkable changes in connectivity across ovulation. This is not surprising (e.g., ROIs were selected based upon evidence for their

contribution to emotional eating) and highlights how GIMME will only detect heterogeneity if it exists.

Analyses also demonstrate how parameter estimates from the person-specific GIMME models (β's) can be linked to hormones and clinical status. These illustrative associations are generally consistent with past work highlighting biological dysregulation in women who engage in emotional eating (Klump et al., 2013; Klump et al., 2014), but findings may change in larger samples with different patterns of heterogeneity. In typical samples, the default mode network has greater activation than the reward network during rest, but during tasks, the reward network has greater activation than the default mode network; this toggling reflects healthy brain function (Fox & Greicius, 2010). Interestingly, cases had weaker connections in the default model and weaker ovarian hormone relations with connections in the default mode than controls, potentially suggesting that hormones contribute to resting state "mis-activation" relevant to emotional eating and anxiety.

GIMME uniquely provided these insights. It is a person-specific approach that takes advantage of time-indexed information to model the *direction* of connections between brain regions (e.g., RLP predicts LLP), going beyond bidirectionality, and it indicates whether that prediction is concurrent or lagged in time. GIMME simultaneously models homogeneity and heterogeneity. It affords population-level generalizations (from group-level connections), while addressing the heterogeneity prevalent in neuroimaging and clinical data that is often ignored and leads to spurious findings (as in the average-level analysis).

There are assumptions, limitations, and nuances to GIMME, so researchers should reference detailed tutorials before implementation (Beltz & Gates, 2017; Lane & Gates, 2017), but there are key considerations for a valid and reliable application. First, because GIMME is a

person-specific approach to the analysis of intra-individual variation, power comes from time series length (not sample size); for fMRI research, the more volumes, the better! The optimal length depends on many factors, such as the number of ROIs, nature of the data, and pattern of estimable effects (Beltz & Gates, 2017). Using procedures similar to those employed here, simulations suggest that GIMME results from time series with 30 observations are less accurate than those with 120 observations (Lane, Gates, Pike, Beltz, & Wright, in press). Second, GIMME models observed data and is fit in a structural equation modeling framework, so researchers should monitor potential problems with convergence and standard errors. Third, GIMME assumes that all time-indexed information is handled within the model (i.e., residuals are white noise). This assumption can be explicitly evaluated and becomes increasingly important with short measurement intervals (Beltz & Molenaar, 2015). Fourth, the direction of prediction for contemporaneous connections can be validated using a version of GIMME (for Multiple Solutions) that generates competing models (Beltz & Molenaar, 2016).

GIMME has enormous potential for the study of clinical phenomena, including eating disorders, because it addresses heterogeneity in etiology, presentation, and treatment response through person-specific modeling of intra-individual variation after detecting commonalities across a sample and *a priori* subgroups, such as twin pairs. GIMME, therefore, complements traditional between-subjects approaches to inter-individual variation while accurately reflecting individual-level processes that can be leveraged in precision healthcare.

References

Amianto, F., D'Agata, F., Lavagnino, L., Caroppo, P., Abbate-Daga, G., Righi, D., . . . Fassino, S. (2013). Intrinsic connectivity networks within cerebellum and beyond in eating disorders. *Cerebellum, 12*(5), 623-631. doi: 10.1007/s12311-013-0471-1

Asarian, L., & Geary, N. (2013). Sex differences in the physiology of eating. *American Journal of Physiology: Regulatory Integrative and Comparative Physiology, 305*(11), R1215-R1267. doi: 10.1152/ajpregu.00446.2012

Becker, J. B. (2009). Sexual differentiation of motivation: A novel mechanism? *Hormones and Behavior, 55*(5), 646-654. doi: 10.1016/j.yhbeh.2009.03.014

Beltz, A. M., & Gates, K. M. (2017). Network mapping with GIMME. *Multivariate Behavioral Research, 52*(6), 789-804. doi: 10.1080/00273171.2017.1373014

Beltz, A. M., Gates, K. M., Engels, A. S., Molenaar, P. C. M., Pulido, C., Turrisi, R., . . . Wilson, S. J. (2013). Changes in alcohol-related brain networks across the first year of college: A prospective pilot study using fMRI effective connectivity mapping. *Addictive Behaviors, 38*(4), 2052-2059. doi: 10.1016/j.addbeh.2012.12.023

Beltz, A. M., & Molenaar, P. C. M. (2015). A posteriori model validation for the temporal order of directed functional connectivity maps. *Frontiers in Neuroscience, 9*. doi: 10.3389/fnins.2015.00304

Beltz, A. M., & Molenaar, P. C. M. (2016). Dealing with multiple solutions in structural vector autoregressive models. *Multivariate Behavioral Research, 51*(2-3), 357-373. doi: 10.1080/00273171.2016.1151333

Birkhoff, G. D. (1931). Proof of the ergodic theorem. *Proceedings of the National Academy of Sciences of the United States of America, 17*, 656-660. doi: 10.1073/pnas.17.12.656

Boehm, I., Geisler, D., King, J. A., Ritschel, F., Seidel, M., Araujo, Y. D., . . . Ehrlich, S. (2014). Increased resting state functional connectivity in the fronto-parietal and default mode network in anorexia nervosa. *Frontiers in Behavioral Neuroscience, 8*. doi: 10.3389/fnbeh.2014.00346

Bohon, C., & Stice, E. (2011). Reward abnormalities among women with full and subthreshold bulimia nervosa: A functional magnetic resonance imaging study. *International Journal of Eating Disorders, 44*(7), 585-595. doi: 10.1002/eat.20869

Borsboom, D., & Cramer, A. O. J. (2013). Network analysis: An integrative approach to the structure of psychopathology. *Annual Review of Clinical Psychology, 9*, 91-121. doi: 10.1146/annurev-clinpsy-050212-185608

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York, NY: Guilford Press.

Cattell, R. B. (1952). The three basic factor-analytic designs: Their interralations and derivatives *Psychological Bulletin*(49), 499-520. doi: doi.ord/10.1037/h0054245

Cha, J., Ide, J. S., Bowman, F. D., Simpson, H. B., Posner, J., & Steinglass, J. E. (2016). Abnormal reward circuitry in anorexia nervosa: A longitudinal, multimodal MRI study. *Human Brain Mapping, 37*(11), 3835-3846. doi: 10.1002/hbm.23279

Cowdrey, F. A., Filippini, N., Park, R. J., Smith, S. M., & McCabe, C. (2014). Increased resting state functional connectivity in the default mode network in recovered anorexia nervosa. *Human Brain Mapping, 35*(2), 483-491. doi: 10.1002/hbm.22202

Davidson, K. W., & Cheung, Y. K. (2017). Envisioning a future for precision health psychology: Innovative applied statistical approaches to N-of-1 studies. *Health Psychology Review, 11*(3), 292-294. doi: 10.1080/17437199.2017.1347514

Favaro, A., Santonastaso, P., Manara, R., Bosello, R., Bommarito, G., Tenconi, E., & Di Salle, F. (2012). Disruption of visuospatial and somatosensory functional connectivity in anorexia nervosa. *Biological Psychiatry, 72*(10), 864-870. doi: 10.1016/j.biopsych.2012.04.025

Forbes, M. K., Wright, A. G. C., Markon, K. E., & Krueger, R. F. (2017). Evidence that psychopathology symptom networks have limited replicability. *Journal of Abnormal Psychology, 126*(7), 969-988. doi: 10.1037/abn0000276

Fox, M. D., & Greicius, M. (2010). Clinical applications of resting state functional connectivity. *Frontiers in Systems Neuroscience, 4*. doi: 10.3389/fnsys.2010.00019

Gambhir, S. S., Ge, T. J., Vermesh, O., & Spitler, R. (2018). Toward achieving precision health. *Science Translational Medicine, 10*(430). doi: 10.1126/scitranslmed.aao3612

García-García, I., Jurado, M. A., Garolera, M., Marqués-Iturria, I., Horstmann, A., Segura, B., . . . Neumann, J. (2015). Functional network centrality in obesity: A resting-state and task fMRI study. *Psychiatry Research-Neuroimaging, 233*(3), 331-338. doi: 10.1016/j.pscychresns.2015.05.017

García-García, I., Jurado, M. A., Garolera, M., Segura, B., Sala-Llonch, R., Marqués-Iturria, I., . . . Junqué, C. (2013). Alterations of the salience network in obesity: A resting-state fMRI study. *Human Brain Mapping, 34*(11), 2786-2797. doi: 10.1002/hbm.22104

Gates, K. M., & Molenaar, P. C. M. (2012). Group search algorithm recovers effective connectivity maps for individuals in homogeneous and heterogeneous samples. *Neuroimage, 63*(1), 310-319. doi: 10.1016/j.neuroimage.2012.06.026

Gates, K. M., Molenaar, P. C. M., Hillary, F. G., Ram, N., & Rovine, M. J. (2010). Automatic search for fMRI connectivity mapping: An alternative to Granger causality testing using

formal equivalences among SEM path modeling, VAR, and unified SEM. *Neuroimage, 50*(3), 1118-1125. doi: 10.1016/j.neuroimage.2009.12.117

Gaudio, S., Piervincenzi, C., Zobel, B. B., Montecchi, F. R., Riva, G., Carducci, F., & Quattrocchi, C. C. (2015). Altered resting state functional connectivity of anterior cingulate cortex in drug naive adolescents at the earliest stages of anorexia nervosa. *Scientific Reports, 5*. doi: 10.1038/srep10818

Gusnard, D. A., & Raichle, M. E. (2001). Searching for a baseline: Functional imaging and the resting human brain. *Nature Reviews Neuroscience, 2*(10), 685-694. doi: 10.1038/35094500

Klump, K. L., Culbert, K. M., & Sisk, C. L. (2017). Sex differences in binge eating: Gonadal hormone effects across development. *Annual Review of Clinical Psychology, 13*, 183-207. doi: 10.1146/annurev-clinpsy-032816-045309

Klump, K. L., Hildebrandt, B. A., O'Connor, S. M., Keel, P. K., Neale, M., Sisk, C. L., . . . Burt, S. A. (2015). Changes in genetic risk for emotional eating across the menstrual cycle: a longitudinal study. *Psychological Medicine, 45*(15), 3227-3237. doi: 10.1017/s0033291715001221

Klump, K. L., Keel, P. K., Racine, S. E., Burt, S. A., Neale, M., Sisk, C. L., . . . Hu, J. Y. (2013). The interactive effects of estrogen and progesterone on changes in emotional eating across the menstrual cycle. *Journal of Abnormal Psychology, 122*(1), 131-137. doi: 10.1037/a0029524

Klump, K. L., Racine, S. E., Hildebrant, B., Burt, A. A., Neale, M., Sisk, C. L., . . . Keel, P. K. (2014). Influences of ovarian hormones on dysregulated eating: A comparison of

associations in women with versus women without binge episodes. *Clinical Psychological Science, 2*(3), 545 - 559. doi: 10.1177/2167702614521794

Kullmann, S., Giel, K. E., Teufel, M., Thiel, A., Zipfel, S., & Preissl, H. (2014). Aberrant network integrity of the inferior frontal cortex in women with anorexia nervosa. *Neuroimage-Clinical, 4*, 615-622. doi: 10.1016/j.nicl.2014.04.002

Lane, S. T., & Gates, K. M. (2017). Automated selection of robust individual-level structural equation models for time series data. *Structural Equation Modeling: A Multidisciplinary Journal, 24*(5), 768-782. doi: 10.1080/10705511.2017.1309978

Lane, S., Gates, K., & Molenaar, P. (2017). gimme: Group iterative multiple model estimation. [Computer software manual]. Retrieved from https://CRAN.R-project.org/package=gimme.

Lane, S. T., Gates, K. M., Pike, H. K., Beltz, A. M., & Wright, A. G. C. (in press). Uncovering general, shared, and unique temporal patterns in ambulatory assessment data. *Psychological Methods*.

Lavagnino, L., Amianto, F., D'Agata, F., Huang, Z. R., Mortara, P., Abbate-Daga, G., . . . Northoff, G. (2014). Reduced resting-state functional connectvity of the somatosensory cortex predicts psychopathological symptoms in women with bulimia nervosa. *Frontiers in Behavioral Neuroscience, 8*. doi: 10.3389/fnbeh.2014.00270

Lee, S., Kim, K. R., Ku, J., Lee, J. H., Namkoong, K., & Jung, Y. C. (2014). Resting-state synchrony between anterior cingulate cortex and precuneus relates to body shape concern in anorexia nervosa and bulimia nervosa. *Psychiatry Research-Neuroimaging, 221*(1), 43-48. doi: 10.1016/j.pscychresns.2013.11.004

Marsh, R., Steinglass, J. E., Gerber, A. J., O'Leary, K. G., Wang, Z., Murphy, D., . . . Peterson, B. S. (2009). Deficient activity in the neural systems that mediate self-regulatory control in bulimia nervosa. *Archives of General Psychiatry, 66*(1), 51-63. doi: 10.1001/archgenpsychiatry.2008.504

McFadden, K. L., Tregellas, J. R., Shott, M. E., & Frank, G. K. W. (2014). Reduced salience and default mode network activity in women with anorexia nervosa. *Journal of Psychiatry & Neuroscience, 39*(3), 178-188. doi: 10.1503/jpn.130046

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. doi: 10.1207/s15366359mea0204_1

Molenaar, P. C. M., & Campbell, C. G. (2009). The new person-specific paradigm in psychology. *Current Directions in Psychological Science, 18*(2), 112-117. doi: 10.1111/j.1467-8721.2009.01619.x

Price, R. B., Lane, S., Gates, K., Kraynak, T. E., Horner, M. S., Thase, M. E., & Siegle, G. J. (2017). Parsing heterogeneity in the brain connectivity of depressed and healthy adults during positive mood. *Biological Psychiatry, 81*(4), 347-357. doi: 10.1016/j.biopsych.2016.06.023

Smith, S. M., Miller, K. L., Salimi-Khorshidi, G., Webster, M., Beckmann, C. F., Nichols, T. E., . . . Woolrich, M. W. (2011). Network modelling methods for FMRI. *Neuroimage, 54*(2), 875-891. doi: 10.1016/j.neuroimage.2010.08.063

Sörbom, D. (1989). Model modification. *Psychometrika, 54*(3), 371-384. doi: 10.1007/bf02294623

Sysko, R., Hildebrandt, T., Wilson, G. T., Wilfley, D. E., & Agras, W. S. (2010). Heterogeneity moderates treatment response among patients with binge eating disorder. *Journal of Consulting and Clinical Psychology, 78*(5), 681-690. doi: 10.1037/a0019735

Van Strien, T., Frijters, J. E., Bergers, G. P., & Defares, P. B. (1986). The Dutch Eating Behavior Questionnaire (DEBQ) for assessment of restrained, emotional, and external eating behavior. *The International Journal of Eating Disorders, 5*(2), 295-315.

Acknowledgments

Author Manuscript

Figure Captions

Figure 1. Study ROIs overlaid on a template brain, with a right anterior cutout. Red ROIs are part of the default mode network, with the right lateral parietal (RLP) and medial prefrontal cortex (MPFC) pictured, and the left lateral parietal (LLP) and posterior cingulate cortex (PCC) not pictured. Green ROIs are part of the reward network, with the right striatum (RS), left striatum (LS), and right orbitofrontal cortex (ROFC) pictured, and the left orbitofrontal cortex (LOFC) not pictured. Yellow ROIs are part of the cognitive control network, with the right dorsolateral prefrontal cortex (RDLPFC) and left dorsolateral prefrontal cortex (LDLPFC) pictured.

Figure 2. Final GIMME models for a twin pair discordant for emotional eating and scanned twice across the menstrual cycle; this is Twin Pair 1 listed in Tables 2 and 3. Ellipses reflect ROIs that constitute the default mode (red), reward (green), and cognitive control (yellow) networks. Solid lines depict contemporaneous connections, and dashed lines depict lagged connections. Thick black lines depict group-level connections (estimated for everyone in the sample), thick gray lines depict twin pair-level connections (estimated for all scans from this twin pair), and thin black lines depict individual-level connections (uniquely estimated for an individual and scan). All connections have associated β weights. The four networks fit the data well; see Table 3.

Figure 3. Final GIMME models for a twin pair discordant for anxiety and scanned twice across the menstrual cycle; this is Twin Pair 3 listed in Tables 2 and 3. Ellipses reflect ROIs that constitute the default mode (red), reward (green), and cognitive control (yellow) networks. Solid lines depict contemporaneous connections, and dashed lines depict lagged connections. Thick black lines depict group-level connections (estimated for everyone in the sample), thick gray

lines depict twin pair-level connections (estimated for all scans from this twin pair), and thin

black lines depict individual-level connections (uniquely estimated for an individual and scan).

All connections have associated β weights. The four networks fit the data well; see Table 3.

Figure 4. Neural network (based on unified structural equation modeling) for sample average-

level data, collapsing across time points and twin pairs. Ellipses reflect ROIs that constitute the

default mode (red), reward (green), and cognitive control (yellow) networks. Solid lines depict

contemporaneous connections, and dashed lines depict lagged connections. All connections have

associated β weights, and the model fit the data well; see statistics in text.