

Running head: GENDER IS PERSON-SPECIFIC

## **Daily Gender and Cognition: A Person-Specific Behavioral Network Analysis**

Adriene M. Beltz\* and Dominic P. Kelly

*Department of Psychology, University of Michigan*

\* Corresponding author

2227 East Hall, 530 Church Street

Ann Arbor, MI 48109

[abeltz@umich.edu](mailto:abeltz@umich.edu)

### **Acknowledgements**

The authors thank all past and present members of the M(SD) Lab at the University of Michigan for their tireless assistance with data collection, management, and processing as well as the participants who made this research possible. A. Beltz was supported by the Jacobs Foundation. D. Kelly was supported by a Katz Fellowship from the Institute for Social Research at the University of Michigan.

### Abstract

Gender is person-specific, and it influences and is influenced by a breadth of multidimensional psychological factors, including cognition. Directionality is important for research on gender and cognition, as debate surrounds, for instance, whether masculine self-concepts precede spatial skills, or whether the reverse is true. In order to provide novel insights into the individualized nature of these relations, a person-specific network approach devised by Peter Molenaar and the first author – group iterative multiple model estimation for multiple solutions (GIMME-MS) – was applied to 75-day intensive longitudinal data on gender self-concept (i.e., femininity-masculinity, instrumentality, and expressivity) and cognition (i.e., mental rotations and verbal recall) from 103 young adults. GIMME-MS estimates individualized networks that contain same-day and next-day directed relations, prioritizing relations common across participants. It is ideal for analyzing behavioral time series with unclear directionality, as it generates multiple solutions from which an optimal one is selected. GIMME-MS revealed notable heterogeneity in the presence, direction, and nature of relations from gender self-concept to cognition (~26% of participants) and vice versa (~21% of participants). Findings are wholly novel in revealing the person-specific nature of gender and its cognitive dynamics, yet somehow, unsurprising given the revolutionary corpus of Peter Molenaar.

*keywords:* GIMME, idiographic, intensive longitudinal design, spatial skills, verbal skills

### **Daily Gender and Cognition: A Person-Specific Behavioral Network Analysis**

Peter Molenaar has had an astounding impact on quantitative psychology. Often taking a biocognitive perspective to development, his work theoretically and mathematically challenged the way averages describe individuals, and arguably catalyzed a revolution of person-specific science (Molenaar, 2004; Molenaar & Beltz, 2020). Peter developed and re-envisioned an astonishing number of analytic approaches to accurately model individuals via intensive longitudinal data, or data with many repeated observations over relatively short time intervals (e.g., hours or days) from the same person. One approach he devised later in his career with Kathleen Gates was group iterative multiple model estimation (GIMME; Gates & Molenaar, 2012). Due to its novel integration of group-level information – without averaging – into individualized brain and behavioral networks, GIMME quickly became a seeming mainstay in the analytic toolkit of neuroimagers and behavioral scientists; it also spawned an ever-growing number of extensions. One such extension – GIMME for multiple solutions (GIMME-MS) – was created by Peter and the first author during her postdoctoral position with him (Beltz & Molenaar, 2016). Though a gender developmentalist by training (and in mentoring the next generation of idiographic scholars, including the second author), this author realized the incredible scientific value of applying a quantitative person-specific lens to any and perhaps all psychological inquiries. In a small attempt to illustrate this and to honor Peter’s legacy, while merging it with the authors’ expertise in gender development, the goal of this paper is to determine the person-specific interplay among daily gender self-concept and cognition.

### **Gender and Cognition: Person-Specific and Directional Links**

Gender has many dimensions. It can be manifest in biology, activities and interests, personal-social attributes, relationships, styles and symbols, values through self-concepts and

beliefs, identity, preferences, and behaviors (Ruble et al., 2006). Cognition is a particularly intriguing personal-social behavior, as it changes across development, training, and context, and is also linked to many other gendered dimensions (Newcombe, 2002). For instance, the sex-role mediation hypothesis (Nash, 1979) espouses that – regardless of sex or gender – masculine gender self-concept (e.g., personality qualities or self-perceptions) positively predicts spatial skills, which show a moderate-to-large gender difference favoring men, and that feminine gender self-concept positively predicts verbal skills, which show a small-to-moderate difference favoring women (reviewed in Beltz et al., 2020). The basic idea is that individuals internalize beliefs surrounding gender roles, act in ways consistent with those roles (i.e., personal characteristics), and eventually, think in ways aligned with those characteristics and roles.

Support for this hypothesis is mixed. There is relatively weak evidence for a link between femininity and verbal skills (but see Reilly et al., 2016), but moderate-to-strong evidence for a link between masculinity and spatial skills (Reilly & Neumann, 2013), especially three-dimensional (3D) mental rotations, potentially because it shows the largest gender difference of all spatial skills (Beltz et al., 2020). It is unclear, however, if the latter link is similar in men and women; there is indication that it is stronger in men than women (Reilly & Neumann, 2013), and that it is stronger in women than in men (Kelly & Beltz, 2022; Signorella & Jamison, 1986). Perhaps this lack of clarity is due, in part, to the heterogeneity of gender. It can develop, look, and mean something different for everyone at different times, so it is likely that relations among its diverse dimensions are *person-specific*. Person-specificity would be missed by the averages derived from traditional between-person analyses, as they likely violate tenets of the ergodic theorem regarding the equivalence of between- and within-person variation; this only holds when psychological processes are homogenous across people and over time (Molenaar, 2004).

Moreover, it is unclear if masculinity or femininity *predicts* cognition, or whether the reverse is true. The former is theoretically posited by the sex-role mediation hypothesis, but the extant literature is cross-sectional and based on self-reports of gender with concurrent assessments of cognition; thus, it cannot statistically support directional inferences. Recent research posited that gendered cognition might instead *predict* masculinity and femininity, arguing from a constructivist perspective that individuals extract meaning from their cognitive skillset, and in turn, identify at least partially with its gendered nature (Kelly & Beltz, 2022). Although this same work reported greater variance explained by statistical models in this novel direction, the study was also cross-sectional, leaving unanswered critical questions about the direction of relations among these gendered dimensions. Fortunately, gender and cognition are beginning to be assessed on relatively short timescales (e.g., momentarily or daily) in intensive longitudinal research (e.g., Mehta & Dementieva, 2017; Sliwinski et al., 2018), creating unique opportunities to fill these critical knowledge gaps: Perhaps gender self-concepts and cognition are not fixed attributes of a person, but also, behaviors that meaningfully vary day-to-day both in degree and in relation to each other.

### **Current Study and Methods**

The goal of this study was to examine the person-specificity and directionality of relations between gender self-concept and cognition utilizing intensive longitudinal data combined and GIMME-MS (Beltz & Molenaar, 2016). A sample average network was also estimated to showcase how inaccurate assumptions of homogeneity can misrepresent heterogeneous individuals. Specifically, data from 103 young adults ( $M_{age}=21.98[3.4]$ ; 77% women; 68% White, 25% Asian, and 7% Black/African American; 93% non-Latinx) who participated in a 75-day intensive longitudinal study (with an average 94% response rate) were

analyzed. This is a short report, but detailed sample characteristics and study methods are reported in other papers (e.g., Beltz et al., 2021; Kelly & Beltz, 2021; Kelly et al., 2020; Weigard et al., 2021); thus, only unique considerations for this paper, which is wholly novel in applying a variant of GIMME to understand daily relations between gender self-concept and cognition, are included below. Additional details are provided in Supplemental Materials.

For 75 consecutive evenings, participants completed an approximately 20-minute survey on any Internet-enabled device. The survey contained three aspects of gender self-concept and two gendered cognitive tests that were then utilized in behavioral network analyses.

*3D mental rotations* (validated in Kelly & Beltz, 2021) indexed daily spatial skills. Participants had 2.5 minutes to complete five items. Each item portrayed a target 3D object made of small blocks portrayed on a 2D plane along with four similar objects as responses. Participants had to choose which two objects were accurate rotations of the target in 3D space, and they received one point for each correct response. Thus, scores could range from 0 to 10 each day.

*Delayed paired verbal recall* (validated in Kelly & Beltz, 2021) indexed daily verbal skills and memory, which shows a small-to-moderate gender difference favoring women (Beltz et al., 2020). At the beginning of each survey, participants were shown five word pairs for two seconds each. Each pair was semantically less related than the previous one (see Kelly & Beltz, 2021); this made the later word pairs each day (e.g., unrelated abstract and concrete nouns, such as “wisdom gravel”) more difficult to remember than the early ones (e.g., commonly-associated adjectives and nouns, such as “dining room”). At the end of the survey, participants were presented with the first word in each pair, and they had to type the second word into a text box. They received one point for each correct response; thus, scores could range from 0 to 5 each day.

*Instrumentality* indexed daily masculine and *expressivity* indexed daily feminine gendered personality qualities, consistent with the most common operationalizations of gender self-concept in investigations of the sex-role mediation hypothesis. These were subscales from the Personal Characteristics Scale (Antill et al., 1993), which was modified for daily responding. Participants indicated on a Likert scale from 1 (*almost never*) to 5 (*almost always*) how often they felt or acted instrumental (e.g., competitive, independent) and expressive (e.g., considerate, patient) that day. Composites were computed by averaging the six items from each subscale each day, with high scores reflecting greater instrumentality and expressivity, respectively.

*Self-perceived femininity-masculinity* indexed daily gender expression. This is a continuum-based measure of gender self-concept from a daily adaptation of the Sex Role Identity Scale (Storms, 1979). Participants indicated on a Likert scale from 1 (*not at all*) to 5 (*extremely*) how masculine and then how feminine their personality, actions/appearance, and in general they were that day. A composite was computed by reverse scoring the three feminine items and averaging them with the three masculine items each day, with low scores reflecting greater femininity and high scores reflecting greater masculinity (e.g., Beltz et al., 2021).

## **GIMME-MS**

The daily composite scores from the five gender self-concept and cognition variables for each participant were submitted to GIMME-MS (Beltz & Molenaar, 2016), which is described in Figure 1. Combined across participants, average composite scores were also submitted separately to GIMME-MS for the sample average analysis. Models were fit using the *ms\_allow* argument in the *gimmeSEM* function of the *gimme* package version 0.7-7 in R (Lane et al., 2021). See Supplemental Materials for the equation underlying GIMME-MS and the R code used in this paper. Briefly, GIMME-MS implements sparse person-specific unified structural equation

models (Gates et al., 2010), which estimate directed relations among variables that are contemporaneous (i.e., at the same time point or day) or lagged of order one (i.e., from one day to the next). These relations are added to null networks in an iterative, data-driven procedure using Lagrange Multiplier tests (i.e., modification indices; Sörbom, 1989) to identify relations that will most improve model fit within a grouping procedure that prioritizes homogeneity (i.e., similarity across participants) if it exists. Specifically, relations that would significantly improve model fit for a criterion of the sample (Gates & Molenaar, 2012) are iteratively added for all individuals, yet estimated uniquely from their own data. After pruning, individual-level relations (reflecting heterogeneity across participants) are iteratively added to the model if they significantly improve fit for an individual until all significant relations are exhausted or until the model fits well (whichever occurs first), according to standard indices: root mean square error of approximation (RMSEA)  $\leq .05$ ; standardized root mean residual (SRMR)  $\leq .05$ ; comparative fit index (CFI)  $\geq .95$ ; and non-normed fit index (NNFI)  $\geq .95$ . The model is then pruned.

Throughout the group- and individual-level iterative procedure, multiple solution paths could be estimated. These paths occur when two (or possibly more) modification indices are equivalent. Typically, the equivalent modification indices correspond to contemporaneous relations between the same two variables, but in opposite directions; thus, multiple solutions often reflect unclear directionality between two variables at the same time point. This is likely to occur in data with large contemporaneous relations (e.g., when inter-measurement intervals are long, as in daily data) because these relations are estimated early in GIMME's search when models are sparse or otherwise lacking information from the lagged relations (i.e., autoregressives for a node or lagged relations between nodes) that would make the modification indices favor one direction or the other for the nodes involved (Beltz & Molenaar, 2016;



Weigard et al., in press). In these cases, GIMME-MS creates bifurcating solution paths, estimating multiple models – each with a relation in one of the equivalent directions before continuing to iterate along each path separately; paths could continue to bifurcate and result in a diverse set of solutions.

Here, the optimal solution was then selected from the set following procedures utilized in simulations and empirical applications (Beltz & Molenaar, 2016; Kelly et al., 2020; Weigard et al., in press): Poor-fitting and implausible solutions were excluded, and the remaining model with the lowest AIC was selected. Largescale simulation results indicate that GIMME-MS provides advantages for directional inferences in daily data when variable first-order autoregressives are .3 or lower. In these data, three of the five variables had average autocorrelations less than .3 (i.e., femininity-masculinity, mental rotations, and verbal recall).

Person-specific networks were then characterized by their key features to facilitate inferences (Beltz & Gates, 2017), focusing on the unique information provided by GIMME-MS: the presence and direction of relations between gender self-concept and cognition as well as the frequency of autoregressions. The former directly results from GIMME-MS, which generates models with the optimal direction of statistical prediction between variables. The latter indirectly follows from GIMME-MS because autoregressions are typically estimated for all variables (even if insignificant and small) in other variants of GIMME (Lane et al., 2019; Weigard et al., 2021).

### **Findings and Person-Specific Inferences**

GIMME-MS generated 198 solutions from the five-variate behavioral time series of the 103 participants. The majority (58.3%) of participants had more than one solution, with an average of 1.92 ( $SD=1.02$ ) individual-level networks per person (ranging from 1 to 6). Interestingly, there were no group-level relations, and thus, no group-level multiple solutions.

This likely reflects extreme heterogeneity in the sample regarding gender self-concept and cognition, and is consistent with expectations about the multidimensionality and person-specificity of gender (Beltz et al., 2021; Ruble et al., 2006). Based on the AIC, unique solutions were selected for all participants, except for 6 individuals, who had equivalent models with contemporaneous relations concerning gender self-concept and cognition in opposite directions. (They are included in subsequent analyses of non-directional, but not directional inferences.) Average fit across participants for the selected models was excellent:  $\chi^2(30.42)=30.91, p=.445$ , RMSEA=.017, SRMR=.081, CFI=.987, and NNFI=1.00.

Illustrative networks for four participants are shown in Figure 2. The man's network in Figure 2A (selected from three solutions) fit well,  $\chi^2(31)=30.72, p=.480$ , RMSEA<.001, SRMR=.078, CFI=1.00, and NNFI=1.00. The pattern of daily relations was generally consistent with the sex-role mediation hypothesis for masculine self-concept and cognition, in that daily femininity-masculinity (which evidenced stability via a positive autoregressive relation) positively predicted mental rotations; there was also a feedback loop between daily instrumentality and expressivity. The man's network in Figure 2B (selected from two solutions) also fit well,  $\chi^2(30)=29.47, p=.493$ , RMSEA<.001, SRMR=.078, CFI=1.00, and NNFI=1.00. The pattern of relations was partially consistent with the sex-role mediation hypothesis, in that there were lagged and positive relations between gender-congruent self-concept and cognition (i.e., verbal recall with expressivity and mental rotations with femininity-masculinity). The masculine relation, however, was in the opposite direction of the hypothesis, with mental rotations predicting masculinity, consistent with the variance explained in the recent cross-sectional report offering a constructivist account of gender self-concept and cognition (Kelly & Beltz, 2022). The participants' networks shown in Figures 2C (only solution;  $\chi^2(28)=28.55, p=.436$ , RMSEA=.016,

SRMR=.066, CFI=.991, and NNFI=.985) and 2D (selected from two solutions;  $\chi^2(29)=28.83$ ,  $p=.474$ , RMSEA<.001, SRMR=.084, CFI=1.00, and NNFI=1.00) had patterns of relations that were generally inconsistent with both the sex-role mediation hypothesis and a constructivist explanation. For instance, the woman in Figure 2C had a positive next-day relation from instrumentality to verbal recall (that evidenced stability via a positive autoregressive relation), which in turn, had a next-day relation to expressivity. Moreover, the woman in Figure 2D had an inverse same-day relation from femininity-masculinity to mental rotations, which in turn, had a positive relation to expressivity. These four participants are representative of the full sample in showcasing the person-specificity of daily relations between gender self-concept and cognition as well as the complex ways in which individualized behavioral networks can support – and counter – broader theories.

Nonetheless, patterns across person-specific networks were examined in an attempt to draw some general inferences. For instance, 76.7% of participants had an autoregression for a gender self-concept variable, but only 22.3% had an autoregression for a cognitive variable, potentially suggesting higher psychological inertia for the former than the latter across days (consistent with greater participant exclusions for low variance in gender self-concept than cognition; see Supplemental Materials). Moreover, average network complexity (i.e., number of connections) was 4.58 ( $SD=2.14$ ), with 68.0% of participants having any relation between cognition and gender self-concept (50.5% had a lagged relation, and 45.6% had a contemporaneous one). Specifically, 50.5% had a relation from gender self-concept to cognition (55.7% lagged and 45.4% contemporaneous), and 42.3% of participants had a relation from cognition to gender self-concept (30.9% lagged and 21.6% contemporaneous). It is important to highlight that these are percentages for relations between *any* gender self-concept node (i.e.,

femininity-masculinity, instrumentality, or expressivity) and *either* cognition node (i.e., mental rotations or verbal recall). The percentage of participants' whose networks contained relations between specific pairs of nodes – broken down by direction and timescale – is shown in Table 1. These percentages further showcase sample heterogeneity, as the greatest percentage of participants who had the same relation was 10.31%.

Indeed, only 25.8% of participants had a relation (or more than one) consistent with the sex-role mediation hypothesis, that is, a positive relation (of any timescale) from mental rotations to instrumentality or femininity-masculinity, or a positive relation from verbal recall to expressivity or femininity (reflected by an inverse relation with femininity-masculinity). Only 20.6% of participants had a relation (or more than one) in the opposite direction, consistent with recent constructivist accounts.

Finally, it is powerful to compare these heterogeneous person-specific networks to a homogeneity-assumed network estimated by running GIMME-MS (without group-level relations) on the average time series across the 103 participants. The single resulting network is shown in Figure 3, and it fits the sample-average data well,  $\chi^2(31)=20.43$ ,  $p=.926$ ,  $RMSEA<.001$ ,  $SRMR=.068$ ,  $CFI=1.00$ , and  $NNFI=1.00$ . Expressivity seems to be a network hub, evincing an autoregression as well as same-day prediction of instrumentality and inverse next-day prediction of mental rotations; verbal recall also has an autoregression. Strikingly, no individual participant has this exact network in relation presence, direction, and timescale (not to mention magnitude).

## Discussion

Despite the multidimensionality of gender, theories concerning the relations among dimensions – and how they manifest in individuals' daily lives – are perplexingly universal. For instance, the sex-role mediation hypothesis (Nash, 1979) purports that masculine self-concepts

(both instrumentality and self-expression) predict male-typed spatial skills, and that feminine self-concepts (both expressivity and self-expression) predict female-typed verbal skills. There is, however, increasing questioning of this universality, as research on directionality shows greater statistical prediction from cognition to gender self-concept, and research on gender highlights individual differences (e.g., Kelly & Beltz, 2022; Reilly et al., 2016; Reilly & Neumann, 2013). Thus, the goal of this study was to estimate the direction and person-specificity of 75-day relations between gender self-concept and cognitive variables in behavioral networks using GIMME-MS. GIMME-MS is an ideal data-driven approach for achieving this goal because it generates a set of individualized networks with same-day and next-day relations in different directions (if there is indication that an opposite-direction relation could produce a competing model with a different fit or inference) from which one with the optimal statistical prediction is selected (Beltz & Molenaar, 2016).

In 103 young adults, GIMME-MS revealed significant heterogeneity in the presence and direction of relations between daily gender self-concept and cognition. This was evidenced in two key ways. First, there were no group-level relations, that is, there were not any relations that were significant for at least 75% of the sample. This is consistent with other daily studies on utilizing GIMME-MS (Kelly et al., 2020), and means that results consisted of a set of 103 person-specific networks. Second, only ~26% and ~21% of participants, respectively, had a gender congruent network relation from self-concept to cognition (consistent with the sex-role mediation hypothesis) or from cognition to self-concept (consistent with a recent constructivist explanation). Although the detection of these directed relations is indirectly influenced by other variables and relations in the networks, these small percentages are still surprising given the relatively small networks (with only 5 nodes). This could reflect true differences between trait-

like (evaluated in cross-sectional studies when participants respond *in general*) and state-like (evaluated in intensive longitudinal studies when participants respond *today*) gender self-concept and cognition. It could also reflect support for the non-ergodicity of gender self-concept and cognition, or that average relations across people do not map onto individualized relations within people, consistent with some of Peter's seminal work (Molenaar, 2004). The fact that no participants had a network matching the sample-average network evocatively supports this.

Indeed, the presence and direction of relations in the behavioral networks was person-specific, neither fully supporting nor contradicting existing expectations about gender self-concept and cognition. Instead, findings suggest that different explanations might apply to different individuals. After replication in other samples and with other gender self-concept and cognition measures, they may even suggest that novel theories are required to explain some participants' patterns of relations. For instance, there were several individuals with relations between gender incongruent aspects of self-concept and cognition, such as the positive relations between masculine instrumentality and verbal recall seen for the women in Figures 2C and 2D. Thus, it will be important for future work to examine whether network features systematically relate to individual differences, including gender (incorporating hormonal milieu), age, and social experiences, as this could guide empirical and theoretical inference-making. Generally, findings highlight the importance of matching a research question to an analytic approach. Indeed, this is a specific example of how a methodological innovation (i.e., GIMME-MS) facilitated a novel substantive application on directional relations between gender self-concept and cognition, with results in turn, highlighting the need for future theory development.

There are key considerations in this research. Concerns common to intensive longitudinal studies apply, including sample representativeness and repeated testing (see Beltz et al., 2021;

Kelly & Beltz, 2021; Kelly et al., 2020; Weigard et al., 2021). Response rates were high for the participants included here (see Supplemental Materials), though, and the cognitive measures were validated (Kelly & Beltz, 2021). Other key considerations include the interpretation of relations among gender self-concept and among cognition nodes, or the so-called within-network relations. Also, most work on the sex-role mediation hypothesis uses only gendered personality qualities (e.g., instrumentality and expressivity), although recent work has considered gender expression, such as feminine-to-masculine self-perceptions (e.g., Beltz et al., 2021). It will be important for future studies to explicitly map the daily overlap among these variables.

Other major considerations concern GIMME or GIMME-MS. One reason for GIMME's increasing popularity is its ability to provide insight into both homogeneity (via group-level relations) and heterogeneity (via individual-level relations); in fact, its prioritization of group-over individual-level relations is largely responsible for its excellent performance in simulations compared to other person-specific approaches (Gates & Molenaar, 2012). Unfortunately, empirical data in which no group-level structure is detected – such as those reported here – do not benefit from this feature of GIMME, as the final models are essentially a set of person-specific networks. This raises questions about network reliability, and thus, whether interpretations of heterogeneity are just wishful misnomers for noise (Hoekstra et al., in press). Noise depends on many factors, including study design, time series length, and measurement; there have even been suggestions to lower GIMME's cutoff from .75 to .50 to facilitate detection of group-level structures and reduce the influence of noise (Gates et al., 2020). Yet, the possibility that some psychological processes are truly heterogeneous remains. Thus, there is potential risk and great reward in the study of person-specific processes, which may be inherently less reliable than group-level processes, which in turn, may or may not actually exist.

Timescale, including assessment and analysis, also has significant implications for interpretations of GIMME results. Regarding assessment, the timescales at which gender self-concept and cognition are thought to meaningfully vary matter. For instance, in this intensive longitudinal study and based on other recent studies (Mehta & Dementieva, 2017; Sliwinski et al., 2018), both variables are expected to fluctuate daily. This daily conceptualization is an assumption, though, as participants' self-reports of gender self-concept concerned the past 24 hours, and their cognitive tests reflected their skills during the 20-minute surveys. Regarding analysis, GIMME implements a discrete-time model with lagged and contemporaneous network relations. These relations can provide novel insights into the studied constructs, as most participants had a gender autoregressive but few had a cognitive autoregressive relation, and as more participants had lagged than contemporaneous relations between gender self-concept and cognition (see Table 1). These findings might be artifacts linked to assessment, or they might indicate greater stability in gender than cognition, suggesting that the constructs are operating on different timescales. Thus, future work could more densely measure cognition than gender and analyze the time series using continuous time approaches to structural vector autoregressive models (of which GIMME is one example; see Chow et al., in press). Such an approach holds significant promise, and it should come as no surprise that Peter Molenaar contributed to this breakthrough.

Lastly, GIMME-MS uses a principled approach to provide insights into statistical prediction, but directional contemporaneous inferences may not always be accurate (see Weigard et al., in press). This may be partly due to model selection. Model plausibility and the AIC were used in this study due to their excellent performance in simulations (Beltz & Molenaar, 2016; Weigard et al., in press) and precedent in the empirical literature (e.g., Kelly et al., 2021). Model



selection can be done in several other ways, too, including through cross-validation, different statistical metrics (e.g., the BIC, which agreed with the AIC in 78.26% of cases for these data), or even features of the fitted networks (e.g., presence of autoregressives, which increase reliability; Lane et al., 2019). This is an exciting area for future work, and a further example of the ways in which empirical investigation can spur methodological development. Broadening this work to include other extensions of GIMME could also be insightful. For example, a new extension of GIMME, called hybrid-GIMME, can estimate contemporaneous bidirectional relations between variables that represent the same construct exogenous to the network (e.g., account for “within-network” relations in the error structure; Luo et al., in press). If this approach were employed, it could indirectly impact the direction of relations for cross-domain variables (e.g., relations between gender self-concept and cognition) that are endogenous within the network. Hybrid-GIMME and GIMME-MS are not yet integrated, though, so this work provides yet another example of how empirical applications also encourage methodological innovation.

## **Conclusion**

Peter Molenaar’s impact on the field of quantitative psychology, and its intersections with biology, neuroscience, statistics, and engineering, among other disciplines, has been staggering. Through the application of the ergodic theory to psychological science – a notion that has seemingly gone from heretical to brilliant in the past few decades – he has helped science increasingly meet modern demand for personalized medicine, individualized intervention, and precision education. By integrating a data-driven approach for person-specific model generation and selection, which Peter and the first author devised (i.e., GIMME-MS), findings from this study strikingly align with these notions of personalization, individualization, and precision: The daily relation between gender self-concept and cognition is unique to each person.

### References

- Antill, J. K., Russell, G., Goodnow, J. J., & Cotton, S. (1993). Measures of children's sex typing in middle childhood. *Australian Journal of Psychology*, *45*(1), 25-33. <https://doi.org/10.1080/00049539308259115>
- Beltz, A. M., & Gates, K. M. (2017). Network mapping with GIMME. *Multivariate Behavioral Research*, *52*(6), 789-804. <https://doi.org/10.1080/00273171.2017.1373014>
- Beltz, A. M., Kelly, D. P., & Berenbaum, S. A. (2020). Sex differences in brain and behavioral development. In J. Rubenstein & P. Rakic (Eds.), *Neural Circuit and Cognitive Development* (2nd ed., pp. 585-638). Academic Press. <https://doi.org/10.1016/b978-0-12-814411-4.00027-5>
- Beltz, A. M., Loviska, A. M., & Weigard, A. (2021). Daily gender expression is associated with psychological adjustment for some people, but mainly men. *Scientific Reports*, *11*(1), 1-10. <https://doi.org/10.1038/s41598-021-88279-4>
- Beltz, A. M., & Molenaar, P. C. (2016). Dealing with multiple solutions in structural vector autoregressive models. *Multivariate Behavioral Research*, *51*(2-3), 357-373. <https://doi.org/10.1080/00273171.2016.1151333>
- Chow, S.-M., Losardo, D., Park, J., & Molenaar, P. C. M. (in press). Continuous-time dynamic models: Connections to structural equation models and other discrete-time models. In R. H. Hoyle (Ed.), *Structural Equation Modeling: Concepts, Issues, and Applications*. Guilford.
- Gates, K. M., Fisher, Z., & Bollen, K. A. (2020). Latent variable GIMME. *Psychological Methods*, *25*(2), 227-242. <https://doi.org/10.1037/met0000229>
- Gates, K. M., Molenaar, P. C., 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. <https://doi.org/10.1016/j.neuroimage.2009.12.117>
- 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. <https://doi.org/10.1016/j.neuroimage.2012.06.026>
- Hoekstra, R. H. A., Epskamp, S., & Borsboom, D. (in press). Heterogeneity in individual network analysis: Reality or illusion? *Multivariate Behavioral Research*. <https://doi.org/10.1080/00273171.2022.2128020>
- Kelly, D. P., & Beltz, A. M. (2021). Capturing fluctuations in gendered cognition with novel intensive longitudinal measures. *Assessment*, *28*(7), 1813-1827. <https://doi.org/10.1177/1073191120952888>
- Kelly, D. P., & Beltz, A. M. (2022). Spatial skills and self-perceived masculinity: Considering college major STEM-ness, directionality, and gender. *Sex Roles*, *87*, 251-266. <https://doi.org/10.1007/s11199-022-01308-y>
- Kelly, D. P., Weigard, A., & Beltz, A. M. (2020). How are you doing? The person-specificity of daily links between neuroticism and physical health. *Journal of Psychosomatic Research*, *137*, 110194. <https://doi.org/10.1016/j.jpsychores.2020.110194>
- Lane, S. T., Gates, K. M., Fisher, Z., Arizmendi, C., Molenaar, P. C., Hallquist, M. N., Pike, H. K., Henry, T., Duffy, K., Luo, L., Beltz, A. M., Wright, A. G., Park, J., & Castro Alvarez, S. (2021). *Group Iterative Multiple Model Estimation*. <https://CRAN.R-project.org/package=gimme>

- Lane, S. T., Gates, K. M., Pike, H. K., Beltz, A. M., & Wright, A. G. C. (2019). Uncovering general, shared, and unique temporal patterns in ambulatory assessment data. *Psychological Methods, 24*(1), 480-488. <https://doi.org/10.1037/met0000192>
- Luo, L., Fisher, Z., Arizmendi, C., Molenaar, P. C. M., Beltz, A., & Gates, K. (in press). Estimating both directed and undirected contemporaneous relations in time series data using hybrid-group iterative multiple model estimation. *Psychological Methods*. doi: <https://doi.org/10.1037/met0000485>
- Mehta, C. M., & Dementieva, Y. (2017). The contextual specificity of gender: Femininity and masculinity in college students' same- and other-gender peer contexts. *Sex Roles, 76*, 604–614. <https://doi.org/10.1007/s11199-016-0632-z>
- Molenaar, P. C. M. (2004). A manifesto on psychology as idiographic science: Bringing the person back into scientific psychology, this time forever. *Measurement: Interdisciplinary Research & Perspectives, 2*(4), 201-218. [https://doi.org/10.1207/s15366359mea0204\\_1](https://doi.org/10.1207/s15366359mea0204_1)
- Molenaar, P. C. M., & Beltz, A. M. (2020). Modeling the individual: Bridging nomothetic and idiographic levels of analysis. In A. G. C. Wright & M. N. Hallquist (Eds.), *Handbook of Research Methods in Clinical Psychology* (pp. 327-336). Cambridge University Press. <https://doi.org/10.1017/9781316995808.031>
- Nash, S. C. (1979). Sex role as a mediator of intellectual functioning. In M. A. Wittig & A. C. Peterson (Eds.), *Sex-Related Differences in Cognitive Functioning* (pp. 263-302). Academic Press.
- Newcombe, N. S. (2002). Spatial cognition. In H. Pashler & D. Medin (Eds.), *Steven's Handbook of Experimental Psychology: Memory and Cognitive Processes* (pp. 113–163). Wiley.
- Reilly, D., & Neumann, D. L. (2013). Gender-role differences in spatial ability: A meta-analytic review. *Sex Roles, 68*(9-10), 521-535. <https://doi.org/10.1007/s11199-013-0269-0>
- Reilly, D., Neumann, D. L., & Andrews, G. (2016). Sex and sex-role differences in specific cognitive abilities. *Intelligence, 54*, 147-158. <https://doi.org/10.1016/j.intell.2015.12.004>
- Ruble, D. N., Martin, C. L., & Berenbaum, S. A. (2006). Gender development. In N. Eisenberg, W. Damon, & R. M. Lerner (Eds.), *Handbook of Child Psychology* (Vol. 3, pp. 858-932). John Wiley & Sons, Inc. <https://doi.org/10.1002/9780470147658.chpsy0314>
- Signorella, M. L., & Jamison, W. (1986). Masculinity, femininity, androgyny, and cognitive performance: A meta-analysis. *Psychological Bulletin, 100*(2), 207. <https://doi.org/10.1037/0033-2909.100.2.207>
- Sliwinski, M. J., Mogle, J. A., Hyun, J., Munz, E., Smyth, J. M., & Lipton, R. B. (2018). Reliability and validity of ambulatory cognitive assessments. *Assessment, 25*(1), 14-30. <https://doi.org/10.1177/1073191116643164>
- Sörbom, D. (1989). Model modification. *Psychometrika, 54*(3), 371-384. <https://doi.org/10.1007/bf02294623>
- Storms, M. D. (1979). Sex role identity and its relationships to sex role attributes and sex role stereotypes. *Journal of Personality and Social Psychology, 37*(10), 1779-1789. <https://doi.org/10.1037//0022-3514.37.10.1779>
- Weigard, A. S., Lane, S., Gates, K., & Beltz, A. (in press). The influence of autoregressive relation strength and search strategy on directionality recovery in group iterative multiple model estimation. *Psychological Methods*. <https://doi.org/10.1037/met0000460>
- Weigard, A. S., Loviska, A. M., & Beltz, A. M. (2021). Little evidence for sex or ovarian hormone influences on affective variability. *Scientific Reports, 11*(1), 20925. <https://doi.org/10.1038/s41598-021-00143-7>

Table 1. *Proportion of Participants with Directional Relations Between Gender Self-Concept and Cognition*

| <b>Relation and Direction</b>                 | <b>Relation Timescale</b>    |                     |
|---|------------------------------|---------------------|
|   | <b>Contemporaneous<br/>%</b> | <b>Lagged<br/>%</b> |
| <i>From Gender Self-Concept to Cognition:</i> |                              |                     |
| F-M to Mental Rotations                       | 6.19                         | 8.25                |
| F-M to Verbal Recall                          | 2.06                         | 10.31               |
| Instrumentality to Mental Rotations           | 4.12                         | 6.19                |
| Instrumentality to Verbal Recall              | 3.09                         | 4.12                |
| Expressivity to Mental Rotations              | 7.22                         | 8.25                |
| Expressivity to Verbal Recall                 | 5.15                         | 6.19                |
| <i>From Cognition to Gender Self-Concept:</i> |                              |                     |
| Mental Rotations to F-M                       | 3.09                         | 9.28                |
| Verbal Recall to F-M                          | 0.00                         | 6.19                |
| Mental Rotations to Instrumentality           | 7.22                         | 7.22                |
| Verbal Recall to Instrumentality              | 4.12                         | 2.06                |
| Mental Rotations to Expressivity              | 4.12                         | 9.28                |
| Verbal Recall to Expressivity                 | 5.15                         | 5.15                |

*Note.* N=97 because 6 participants were excluded for having equivalent best-fitting networks with opposite inferences about gender self-concept and cognition relation directionality. Gender self-concept variables are femininity-to-masculinity (F-M), instrumentality, and expressivity; cognition variables are mental rotations and verbal recall.

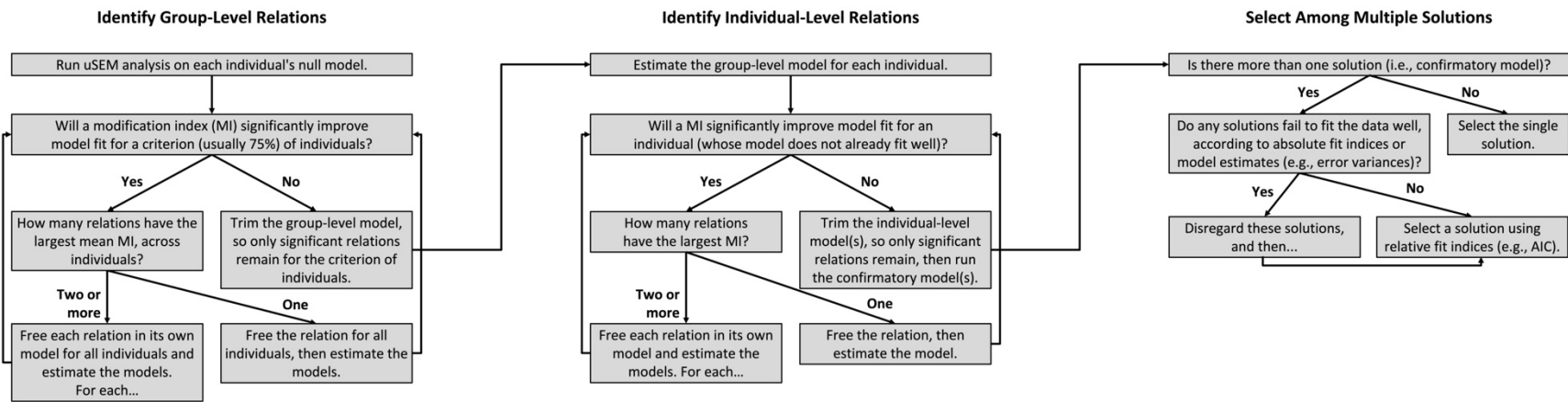
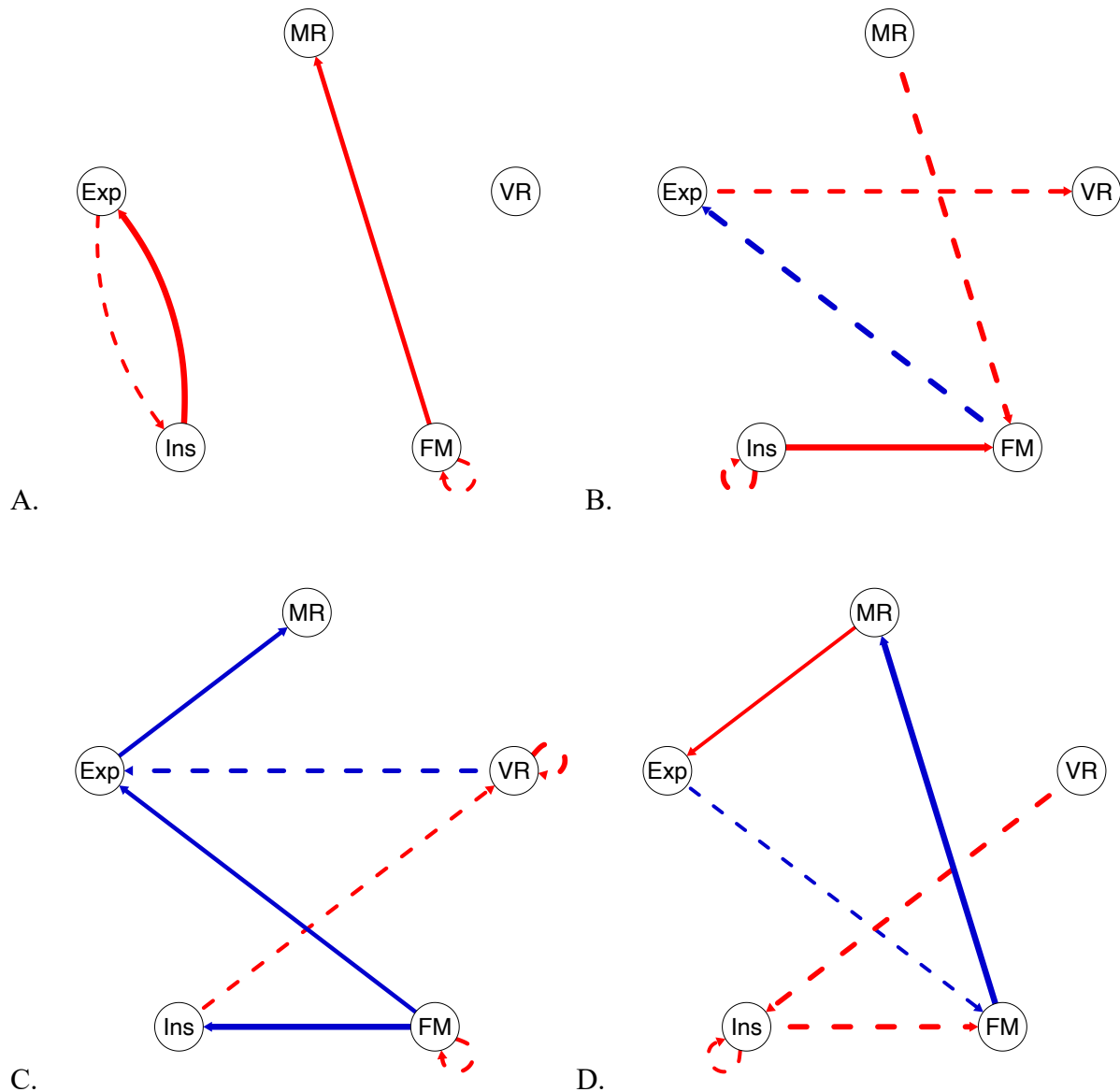
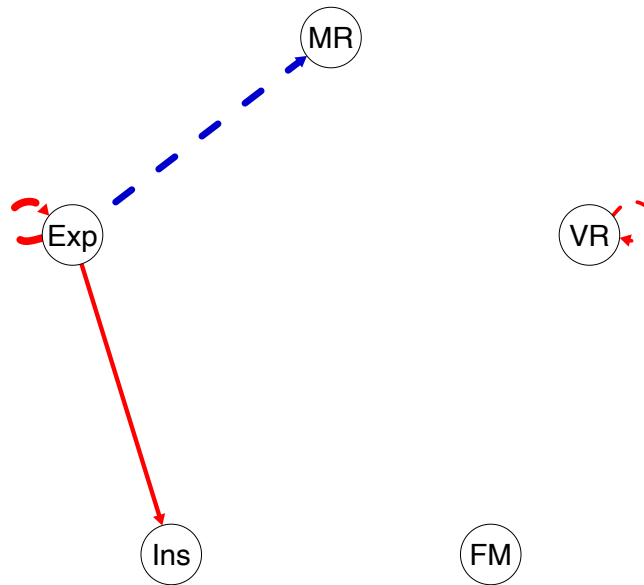


Figure 1. Schematic of group iterative multiple model estimation for multiple solutions (GIMME-MS; Beltz & Molenaar, 2016) implemented in this study. GIMME-MS creates person-specific behavioral networks that reflect directed first order lagged and contemporaneous relations among variables (i.e., unified structural equation models, uSEMs, Gates et al., 2010) at the group- and individual-level. Multiple solutions are generated when Lagrange Multiple tests (i.e., modification indices, Sörbom, 1989) are equal for two or more relations in an iteration; this can occur at the group- or individual-level. If multiple solutions are estimated, then one can be selected after excluding poor-fitting models (see criteria in main text) by comparing relative model fits (e.g., using the Akaike Information Criterion, AIC). Modified with permission from Beltz and Molenaar (2016); see <https://www.tandfonline.com/>.



*Figure 2.* Illustrative behavioral networks for four participants resulting from GIMME-MS, illustrating the person-specificity of the presence, direction, and magnitude of daily relations between gender self-concept (FM, Ins, and Exp) and cognition (MR and VR). Solid lines are contemporaneous (same-day) relations, dashed lines are first order lagged (next-day) relations, red lines are positive relations, blue lines are inverse relations, and line thickness reflects relation magnitude. All models fit well. MR: Mental rotations; VR: Verbal recall; FM: Femininity-to-masculinity; Ins: Instrumentality; E: Expressivity.



*Figure 3.* Behavioral network resulting from GIMME-MS for the sample average data, illustrating the ways in which heterogeneous individual-level processes may be obscured when homogeneity is falsely assumed. The network reflects daily relations between gender self-concept (FM, Ins, and Exp) and cognition (MR and VR) in the sample-averaged time series, and no person-specific network had this exact pattern of relations. Solid lines are contemporaneous (same-day) relations, dashed lines are first order lagged (next-day) relations, red lines are positive relations, blue lines are inverse relations, and line thickness reflects relation magnitude. The model fits well. MR: Mental rotations; VR: Verbal recall; FM: Femininity-to-masculinity; Ins: Instrumentality; E: Expressivity.

**Supplemental Materials for**  
**“Daily Gender and Cognition: A Person-Specific Behavioral Network Analysis”**

**Participants**

Study methods were approved by the Institutional Review Board – Health Sciences and Behavioral Sciences at the University of Michigan (HUM# 00119536). All participants provided written informed consent. Out of 175 participants who completed the 75-day intensive longitudinal study, 103 were included in this report. There was a greater number of women than men due to other aims of the parent intensive longitudinal study (e.g., related to oral contraceptive use). Reasons for exclusion were: starting the intensive longitudinal study before the mental rotations test was included ( $n=15$ ), having less than an 80% completion rate for the intensive longitudinal study or a cognitive measure within it, raising concerns about data fidelity and model accuracy due to missingness ( $n=43$ ), having little-to-no variation in a gender self-concept measure ( $n=9$ ), and not having a plausible model estimated by GIMME-MS ( $n=5$ ).

**Procedures**

Participants were recruited from a university community. They completed a 60-minute monitored intake survey via Qualtrics in a research laboratory. In the intake survey, they provided demographic information and completed several cognitive tests and psychological questionnaires. Next, they completed a 75-day intensive longitudinal study. Every evening at 5PM, they received a personalized link to a Qualtrics survey that took approximately 20 minutes to complete. The surveys could be completed on any Internet-enabled device, and they contained the tests and questionnaires reported in this paper, among others. Participants received \$1 for each daily survey they completed, but this doubled to \$2 for all surveys if they completed at least



80% of them; participants also received a \$50 bonus if they completed at least 90% of the daily surveys.

### GIMME-MS

The equation that underlies both GIMME and GIMME-MS is below. GIMME-MS just differs from GIMME in the procedures used to fit the model. Specifically, during the iterative data-driven addition of relations to each individual's network, relations that would equivalently and maximally improve model fit are added to different solutions paths that then independently iterate to generate a set of solutions from which an optimal one can be selected.

$$\eta_i(t) = (\mathbf{A}_i + \mathbf{A}^g)\eta_i(t) + (\mathbf{\Phi}_{1,i} + \mathbf{\Phi}_1^g)\eta_i(t-1) + \zeta_i(t),$$

in which  $\eta$  represents the 5-variate, up to 75-day time series ( $t$ ) to be explained for each of 103 individuals ( $i$ ). The timeseries is explained by a 5x5 matrix (with a fixed diagonal) of contemporaneous relations ( $\mathbf{A}$ ), some of which are estimated for everyone in the sample/group ( $g$ ) and some of which are estimated only for an individual. It is also explained by a 5x5 matrix (with autoregressives on the diagonal) of first order lagged relations ( $\mathbf{\Phi}$ ), some of which are estimated for everyone in the group and some of which are estimated only for an individual. Residuals are reflected in the 5-variate  $\zeta$  errors with a zero mean, diagonal covariance matrix, and assumed lack of temporal dependencies.

The R code used to estimate the person-specific behavioral networks in GIMME-MS is below. Arguments related to GIMME-MS are *ms\_allow* and *ms\_tol*. See the *gimme* page in Cran for details (<https://cran.r-project.org/web/packages/gimme>).

```
install.packages("gimme")
library(gimme)
MBRmodel <- gimmeSEM(data = "Input",
                     out = "Output",
                     sep = "",
                     header = TRUE,
```

```
ar = FALSE,  
plot = TRUE,  
subgroup = FALSE,  
sub_feature = "lag & contemp",  
#sub_method = "Walktrap",  
confirm_subgroup = NULL,  
exogenous = NULL,  
paths = NULL,  
conv_vars = NULL,  
conv_length = 16,  
conv_interval = 1,  
mult_vars = NULL,  
mean_center_mult = FALSE,  
standardize = TRUE,  
groupcutoff = .75,  
subcutoff = .5,  
diagnos = FALSE,  
ms_allow = TRUE,  
ms_tol = 1e-5,  
lv_model = NULL,  
lv_estimator = "miiv",  
lv_scores = "regression",  
lv_miiv_scaling = "first.indicator",  
lv_final_estimator = "miiv")
```