Behaviour Research and Therapy 180 (2024) 104574

Contents lists available at ScienceDirect

Behaviour Research and Therapy

journal homepage: www.elsevier.com/locate/brat

Heterogeneity in suicide risk: Evidence from personalized dynamic models

Daniel D.L. Coppersmith [a,*], Evan M. Kleiman [b], Alexander J. Millner [a,c], Shirley B. Wang [a], Cara Arizmendi [d], Kate H. Bentley [a,e], Dylan DeMarco [a], Rebecca G. Fortgang [a,e], Kelly L. Zuromski [a], Joseph S. Maimone [f], Adam Haim [g], Jukka-Pekka Onnela [h], Suzanne A. Bird [e], Jordan W. Smoller [e], Patrick Mair [a], Matthew K. Nock [a,c,e]

[a] Harvard University, Department of Psychology, USA
[b] Rutgers, The State University of New Jersey, Department of Psychology, USA
[c] Franciscan Children's, Mental Health Research, USA
[d] Duke University School of Medicine, Department of Population Health Sciences, USA
[e] Massachusetts General Hospital, Department of Psychiatry, USA
[f] Boston University, Department of Psychology, USA
[g] National Institute of Mental Health, USA
[h] Harvard T. H. Chan School of Public Health, Department of Biostatistics, USA

ARTICLE INFO

Keywords:
Suicide
Suicidal ideation
Heterogeneity
Hopelessness
Ecological momentary assessment

ABSTRACT

Most theories of suicide propose within-person changes in psychological states cause suicidal thoughts/behaviors; however, most studies use between-person analyses. Thus, there are little empirical data exploring current theories in the way they are hypothesized to occur. We used a form of statistical modeling called group iterative multiple model estimation (GIMME) to explore one theory of suicide: The Interpersonal Theory of Suicide (IPTS). GIMME estimates personalized statistical models for each individual and associations shared across individuals. Data were from a real-time monitoring study of individuals with a history of suicidal thoughts/behavior (adult sample: participants = 111, observations = 25,242; adolescent sample: participants = 145, observations = 26,182). Across both samples, none of theorized IPTS effects (i.e., contemporaneous effect from hopeless to suicidal thinking) were shared at the group level. There was significant heterogeneity in the personalized models, suggesting there are different pathways through which different people come to experience suicidal thoughts/behaviors. These findings highlight the complexity of suicide risk and the need for more personalized approaches to assessment and prediction.

For centuries, scholars have pondered why people decide to intentionally end their lives. Suicide challenges the idea that all animals possess an instinctual drive to survive, making it one of the most perplexing aspects of human nature (Minois, 2001). Psychological scientists have proposed theories of the causes of suicidal thoughts and behaviors (Millner, Robinaugh, & Nock, 2020; Selby, Jr, & Ribeiro, 2014).

Most theories of the causes of suicide propose that there are psychological processes that change within individuals (e.g., increases in hopelessness) and cause suicidal thoughts and behaviors to occur. However, theories have been almost exclusively tested at the aggregate/group level (i.e., between people). Analyses conducted this way make an underlying assumption that there is a singular psychological pathway leading to suicide. However, the causes of suicide may differ from person to person (Kaurin, Dombrovski, Hallquist, & Wright, 2021; Millner et al., 2020; Sewall & Wright, 2021) and group-level results cannot be assumed to generalize to all individuals (Barlow & Nock, 2009; Fisher, Medaglia, & Jeronimus, 2018; Hamaker, 2012; Molenaar, 2004). One reason for this lack of generalizability is the issue of ergodicity (Fisher et al., 2018; Molenaar, 2004)). For results to translate from the group to the individual (i.e., ergodicity), between-people variability (group) must be equivalent to within-person variability (individual). Past research suggests this is not the case in human subjects research (Fisher et al., 2018) and data is nonergodic.

Fortunately, real time monitoring methods allow us to examine changes in within-person processes that lead up to suicidal thoughts and behaviors. Real-time monitoring refers to data collected in an individual's natural environment, intensively and repeatedly over time (Stone, Schneider, & Smyth, 2023; Wright & Zimmermann, 2019).

* Corresponding author. Harvard University, 33 Kirkland Street, Cambridge, MA, 02138, USA.
E-mail address: dcoppersmith@g.harvard.edu (D.D.L. Coppersmith).

https://doi.org/10.1016/j.brat.2024.104574
Received 24 August 2022; Received in revised form 9 May 2024; Accepted 22 May 2024
Available online 23 May 2024

Studies have begun to use real-time monitoring to examine theories of suicide (Gee, Han, Benassi, & Batterham, 2020; Kleiman, Glenn, & Liu, 2023). These studies, however, have used statistical models that rely on the average effects across people (Czyz, Horwitz, Arango, & King, 2019; Hallensleben et al., 2019; Kleiman et al., 2017). Thus, there is still little information about individual-level processes (Kaurin et al., 2021).

One of most prominent current theories of suicide is the Interpersonal Theory of Suicide (IPTS) (Joiner, 2007; Van Orden et al., 2010). The IPTS makes predictions about both suicidal ideation and suicidal behavior. The theory proposes that suicidal desires (e.g., "I want to kill myself") result from a combination of perceived burdensomeness (e.g., "I am a burden"), thwarted belongingness (e.g., "I am alone"), and hopelessness about that perceived burdensomeness and thwarted belongingness. The pathways hypothesized in the original theory (Van Orden et al., 2010) and reiterated in a meta-analysis of the theory (Chu et al., 2017) are depicted in Fig. 1. Specifically, perceived burdensomeness and thwarted belongingness are proposed to be proximal predictors of hopelessness and hopelessness is a proximal predictor of suicidal desire. The IPTS asserts that perceived burdensomeness, thwarted belongingness, and hopelessness together are proximal and sufficient causes of suicidal desire.

The IPTS is a prominent and influential theory in the field of suicide research (Hjelmeland & Loa Knizek, 2020; Joiner et al., 2021). It also is proposed to have strong generality. Joiner suggests that "the Interpersonal Theory of Suicide is a universal theory of human suicidal behavior, in the sense that it is proposed to be explanatory regarding suicidal behavior across all spectra of diversity (Joiner et al., 2021, p. 4)," noting that it provides an explanation "for all suicides at all times in all cultures across all conditions" (Dokoupil, 2013). In a meta-analysis of the IPTS including over 100 studies (Chu et al., 2017). The majority (92.3%) of research on the IPTS has been cross-sectional (Chu et al., 2017). To date only a small number of real-time monitoring studies have examined the IPTS (Hallensleben et al., 2019; Kleiman et al., 2017) and all have used statistical models that rely on the average effects across people. Therefore, the designs of most of the research on the IPTS is not well suited to explore the theory because the theory makes proximate within-person predictions about suicidal ideation, but the theory has been largely tested with between-person data on suicidal ideation.

Recent advances in statistical software (Lane, Gates, Pike, Beltz, & Wright, 2019) allow researchers to apply a data-driven approach to within-person data to determine whether some relationships are present for most individuals across the entire sample as opposed to only among some individuals (or not at all). This powerful approach is called group iterative multiple model estimation (GIMME; Beltz & Gates, 2017; Gates & Molenaar, 2012). One benefit of this approach is that it allows researchers to quantify the generality of theoretical models. Rather than merely supporting or falsifying a theory with a single statistical test as is traditionally done (Christensen, Batterham, Soubelet, & Mackinnon, 2013), GIMME allows researchers to estimate subgroups of individuals to whom a theory may apply. In the current study, we leveraged real-time monitoring data and GIMME to explore the generality of the theorized relationships across people in samples at elevated risk for suicidal thoughts and behaviors.

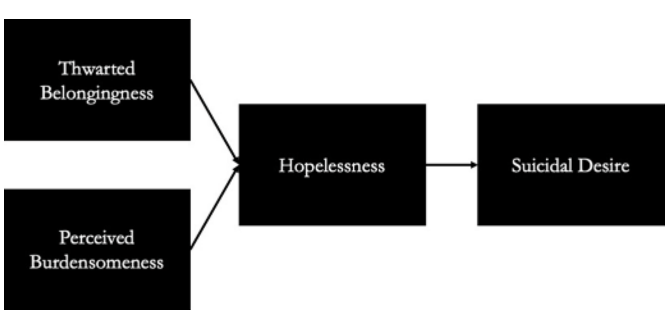

Fig. 1. Pathways theorized by the interpersonal theory of suicide.

1. Current study

The overall aims of the current analysis were to examine the generality and heterogeneity of within-person models of suicidal thinking by testing the IPTS - one of most prominent current theories of suicide. The first aim was to examine the generality of the IPTS. We operationalized the IPTS in the current study through examining bivariate relationships among core constructs. For the IPTS, we specifically would expect most participants to show relationships between perceived burdensomeness and thwarted belongingness with hopelessness and hopelessness with suicidal thinking. The presence of group paths for these relationships would support the IPTS. Also of secondary interest are the relationships between perceived burdensomeness and thwarted belongingness with suicidal thinking. The second aim was to examine the heterogeneity of individual models of the IPTS. We specifically used S-GIMME, a subgrouping algorithm that identifies subgroups of individuals with similar model patterns, for this aim. Through subgrouping and examining individual models, we sought to better understand what percentage of individuals show the presence of paths consistent with the interpersonal theory. The final aim was to compare if statistical approaches, which rely on the average effects across people, result in different support for or against the IPTS than GIMME which relies on individual level effects.

2. Method

2.1. Participants

Participants were from an intensive longitudinal study focused on understanding the natural occurrence of suicidal thoughts and behaviors in the period after hospitalization. Data were downloaded in November 2022. Demographic data are reported in Table 1. In the overall protocol both adults and adolescents were enrolled in the study. For the current analysis, we decided to analyze these two samples separately because of differences in risk factors across age groups (Nock et al., 2008), different clinical recruitment sites of the two samples, and calls for adolescent specific theories of suicide (Hausmann-Stabile, Glenn, & Kandlur, 2021;

Table 1
Sample demographic information.

Characteristic	Adult Sample (n = 111)	Adolescent Sample (n = 145)
Age (years)	32.66 (SD = 11.70)	14.96 (SD = 1.56)
Gender Identity		
Female	57	89
Male	39	17
Nonbinary/Gender nonconforming	4	26
Transgender man	4	6
Transgender woman	2	-
Other	1	4
Missing	4	3
Birth Sex		
Female	67	127
Male	40	15
Missing	4	3
Race		
White	79	106
Black	12	6
Asian	1	12
Pacific Islander	–	–
Native American	–	–
Multiracial	6	12
Other	8	6
Missing	5	3
Ethnicity		
Hispanic/Latino/Latina	22	12

Miller & Prinstein, 2019).

Participants in the adult sample were recruited in the U.S. at a large urban hospital emergency department (adult participants, $n = 111$). The adolescent sample were recruited from an adolescent inpatient psychiatry unit (adolescent participants, $n = 145$). For both groups, inclusion criteria included (a) presenting to the hospital with suicidal thoughts or behaviors and (b) owning a smartphone. For minors, participation also required parental permission. Exclusion criteria for all participants included any factor that impaired their comprehension and participation in the study, including: (1) an inability to speak or write English fluently, (2) the presence of gross cognitive impairment due to florid psychosis, (3) intellectual disability, (4) dementia, (5) acute intoxication, or (6) the presence of extremely agitated or violent behavior. Decisions about these inclusion/exclusion criteria was made by clinicians at the respective hospitals. Prior research suggests that at least 60 observations per participants and at least 10 participants are required to be able to reliably use GIMME (Beltz & Gates, 2017; Lane et al., 2019). Therefore, only participants with at least 60 observations were included in the analysis. In the adult sample, of the original 224 participants who enrolled in the study and completed at least one EMA survey at the time of the data download, 99 were excluded for having fewer than 60 observations and 9 for having no variance in at least one of the variables under examination. Five participants were excluded because GIMME could not fit an adequate model to these participants' data (e.g., singular fit). In the adolescent sample, of the original 253 participants who enrolled in the study and completed at least one EMA survey at the time of the data download, 95 were excluded for having fewer than 60 observations and 13 for having no variance in at least one of the variables under examination.

A concern with excluding 43% (adolescent) to 50% (adult) of the sample is that this may bias the results and that there could be differences between those included and excluded from the analysis. For example, participants excluded may have more severe histories of suicidal thoughts and behaviors. That is to say there could be bias in the type of participants included in the analysis. To test this, we compared participants included and excluded on demographics and baseline suicidal thoughts and behaviors histories for participants who completed the baseline survey and at least one EMA survey. It is important to note that comparing participants on demographic and baseline characteristics cannot speak to potential bias in dynamic within-person effects. For demographics, we compared participants on age, sex assigned at birth, and ethnicity. For suicidal thoughts and behaviors histories, we compared participants on lifetime presence of a suicide attempt, past year presence of a suicide attempt, and number of days with thoughts of killing oneself in the past week. All these variables were measured with a self-report version of the Self-Injurious Thoughts and Behaviors Interview (Fox et al., 2020; Nock, Holmberg, Photos, & Michel, 2007).

For continuous variables (e.g., age) differences in included/excluded were compared with t-tests. For categorical variables (e.g., lifetime presence of a suicide attempt) differences were compared with chi-squared tests. In the adult sample, included/excluded participants did not differ on the examined demographic variables (p values all greater than 0.05). In the adult sample, participants did not differ on lifetime history of a suicide attempt, past year history of a suicide attempt, or number of days with thoughts of killing oneself in the past week (p values all greater than 0.05). In the adolescent sample, included/excluded participants did not differ on the examined demographic variables (p values all greater than 0.05). In the adolescent sample, participants did not differ on lifetime presence of a suicide attempt, past year presence of a suicide attempt, or number of days with thoughts of killing oneself in the past week (p values all greater than 0.05).

2.2. Procedure

The overall study lasted six months. For the first three months, participants were sent smartphone-based momentary assessment surveys six times per day at random times within pre-defined windows using LifeData. For the last three months, participants were sent smartphone-based daily diary assessment surveys once per day. Only data from the first three months were used in the current analysis. Morning and evening surveys were sent at fixed times and daytime surveys were sent randomly within a pre-defined window of time. All procedures were approved by the governing university/hospital institutional review boards.

2.3. Measures

In each smartphone-based survey, participants rated their suicidal thoughts and emotions in the moment (i.e., "right now"). The exact prompt, definition, scale, and anchors of each item used in the analysis are provided in Table 2. We operationalized suicidal thinking with one item on the urge to kill oneself. This type of item has been used in other studies (Bentley et al., 2021; Coppersmith, Millgram, et al., 2023; Wang et al., 2021) and shown predictive validity (Wang et al., 2021). This item was specifically used in the analysis because the IPTS theorizes that perceived burdensomeness, thwarted belongingness, and hopelessness leads to a desire for suicide (e.g., want/desire to kill oneself). The interpersonal items used in this study were "hopeless," "burdensome," and "isolated." For the purposes of this analysis, the isolated item is used to represent the IPTS construct of thwarted belongingness. Loneliness/social isolation has been proposed as a component of thwarted belongingness (Van Orden et al., 2010). While the IPTS posits that hopelessness specifically regarding perceived burdensomeness and thwarted belongingness leads to a desire for suicide (Van Orden et al., 2010), the largest meta-analysis on the IPTS focused on general hopelessness (i.e., not just hopelessness about burdensomeness and thwarted belongingness) (Chu et al., 2017). Consistent with this meta-analysis and other studies (Chu et al., 2020), we used a measure of general hopelessness.

2.4. Analytic approach

All analyses were conducted with GIMME (Gates, Lane, Varangis, Giovanello, & Guskiewicz, 2017). GIMME is a model search algorithm that makes use of the unified structural equation (uSEM) framework to describe temporal processes (Gates & Molenaar, 2012). The details of the algorithm (Gates & Molenaar, 2012), simulation studies (Lane et al., 2019), and tutorials on GIMME have been published elsewhere (Arizmendi, Gates, Fredrickson, & Wright, 2020). Therefore, we briefly described what GIMME consists of, what it produces, and how it compares to other modeling frameworks.

As previously stated, GIMME uses a unified structural equation (uSEM) framework (Kim, Zhu, Chang, Bentler, & Ernst, 2007), combining vector autoregression and structural equation modeling.

Table 2
Real-time measurement items.

Construct	Prompt	Definition	Scale	Anchors
Hopeless	Right now, how much do you feel: Hopeless	"When things are bad you feel like things will never get better."	0 to 10	not at all very much
Burdensome	Right now, how much do you feel: Burdensome	"You, your presence, or things you've done cause other people to suffer hardship."	0 to 10	not at all very much
Isolated	Right now, how much do you feel: Isolated from others	N/A	0 to 10	not at all very much
Suicidal Urge	Right now, how strong is your: Urge to kill yourself	N/A	0 to 10	not at all very much

Vector autoregression is used to model lagged effects, which are the relationships between variables at one assessment and the next assessment. For example, using hopelessness at one assessment to predict suicidal thinking at the next assessment would be a lagged relationship. Structural equation modeling is used to model contemporaneous effects, which are the relationships between variables within the same assessment. As part of the model search procedure, GIMME provides the direction of effects based on significance of effects and model fit indices. In the context of GIMME, direction is defined as one variable predicting another variable after accounting for the influence of the other variables in the model (Wright et al., 2019). For example, hopelessness predicting suicidal thinking when accounting for the effect of suicidal thinking predicting hopelessness. An important limitation regarding directionality within GIMME is that the identification of directionality for contemporaneous paths is not always reliable (Beltz & Molenaar, 2016) and past simulation work suggest imperfect recall and precision of directionality (Lane et al., 2019). This can be especially challenging in the context of data with lower autoregressive relations, such as daily diary or EMA data (Weigard, Lane, Gates, & Beltz, 2023).

There are analytic strategies for addressing this challenge (Weigard et al., 2023) and we used one of them, estimating the autoregressive paths by default, in the current study. Nonetheless, caution is warranted when interpreting the direction of the paths generated by GIMME.

GIMME uses a multi-step procedure for model selection and fitting (Gates et al., 2017; Lane et al., 2019). In the first step, the algorithm searches for paths (both contemporaneous and lagged) that would significantly improve model fit for the majority of individuals in the sample. Following prior studies, we operationalized the majority as 75% of the sample (Gates et al., 2017; Wright et al., 2019). Model fit refers to the ability of the model to characterize the variance and covariance in the data. Significance is operationalized by testing modification indices for each path. It is important to note that during this group-search, the analyses are done for each individual separately rather than any type of pooling or aggregation across individuals.

GIMME then uses the group level paths as priors to search for individual-level paths. Starting with information shared at the group-level to inform individual-level searches has been found to improve the recovery and reliability of individual-level paths (Gates & Molenaar, 2012). GIMME iteratively searches for paths for each individual in the sample. All possible associations have the potential to be modeled but are not included if they are nonsignificant and are detrimental to model fit. Model search continues until an excellent model fit is found for each individual.

There is the additional option of running a subgrouping algorithm to identify subgroups of individuals with similar model patterns. This algorithm, S-GIMME, extends GIMME by identifying subgroups based on shared characteristics of individuals' temporal processes (Lane et al., 2019). In S-GIMME, the search for subgroups occurs in between the group search and the individual search. The purpose of the subgroup search is to refine the model search and to identify subsets of individuals. It specifically creates an adjacency matrix of shared paths across individuals and then applies a community detection algorithm procedure called Walktrap to the matrix. For a more detailed explanation of the estimation procedures see (Gates et al., 2017; Lane et al., 2019). The threshold for a subgroup path is 50%, that is for a path to be considered a subgroup level path 50% of the individuals within that subgroup must contain that path in their model.

GIMME produces group-level paths, subgroup paths (when S-GIMME is used), and individual paths. Group level paths are paths that are present for at least 75% of the sample. Subgroup paths are shared among a subgroup of the sample. Individual paths are paths that are estimated for each individual but do not meet the threshold for group or subgroup level paths. At the individual level, GIMME provides path estimates and model fit information for each individual.

GIMME is unique from other popular analytic approaches to intensive longitudinal data like the real-time monitoring data from this study in several important ways. One of the most common approaches is multilevel modeling (Hamaker & Wichers, 2017). GIMME is unique from multilevel modeling in that multilevel modeling pulls all individual effects into a distribution of effects (Wright et al., 2019). Multilevel models share information across individuals to pool effects (Gelman, Hill, & Yajima, 2012). This makes strong assumptions about the nature of individual effects, can be influenced by extreme cases, and can lead to biases in estimates when there is heterogeneity in the sample (Liu, 2017). GIMME provides no restrictions on individual effects, ensures that all individuals are estimated separately, and each individual has an equal contribution to group-level effects. Another analytic approach is idiographic (i.e., N of 1) structural equation modeling (Fisher, 2015) or vector autoregression (Fisher, Reeves, Lawyer, Medaglia, & Rubel, 2017; Liu, 2017). A limitation of these approaches is scalability (Wright et al., 2019). Traditionally, researchers had to do model fitting for each individual (Fisher et al., 2017). GIMME automates this individual-level model fitting.

All analyses were conducted in R version 4.1.1 with the R package *gimme* (Lane et al., 2021). We included the urge kill oneself item and the individual interpersonal items. We fit the model with the default settings given that these settings have been tested in simulation studies (Gates et al., 2017; Lane et al., 2019) and shown to reliably recover true effects. These settings include defining group-level paths as those that improved model fit in at least 75% of the individual models. We also include all autoregressive paths at the group level, the default for GIMME. This means that all autoregressive paths (e.g., burdensomeness at t predicting burdensomeness at t+1) were estimated for all individuals as opposed to being left open for estimation in the model.

We also used the R package *perturbR* to assess the robustness of subgroups identified (Gates et al., 2019). For this analysis, we took the adjacency matrix, which contains all of the shared paths across participants identified by S-GIMME. Noise is then added to the matrix in increasing amounts to test how robust the matrix is to perturbations. The subgroups identified by S-GIMME on the original adjacency matrix are then compared to the subgroups from the perturbed matrix. If the subgroups are stable, small perturbations to matrix (e.g., slight changes in the paths) should not alter the subgroups. The outcome used for this analysis was the percent of perturbation to edges in the matrix that would result in 20% of subgroup assignments being swapped. A general threshold for stable subgroups is 20% perturbation before 20% of subgroup assignments are swapped (Gates et al., 2019; Groen et al., 2022).

To account for the uneven spacing of the observations in time, we also included two temporal variables in the models. We included the total time since the first observation (in hours) and length of time between observations (in hours). These temporal variables were specified as exogenous variables, which meant that they could predict other variables in the model, but not be predicted by other variables (Arizmendi et al., 2020). The total time since the first observation variable linearly detrends the data in a single model. This is a useful step because an assumption of GIMME is stationarity (i.e., the mean and variance do not change over time). Past work has found similar findings when de-trending the data prior to running GIMME and de-trending with an exogenous temporal node (Webb, Murray, Tierney, & Gates, 2023). An advantage of de-trending within GIMME is that it allows one to quantify the percentage of participants with temporal trends in the variables of interest (Arizmendi et al., 2020). The length of time between observations variable represents the extent to which the time since the last observation has an effect on a variable. For example, the length of time since the last observation predicts hopelessness. This approach to accounting for uneven spacing of observations, which is the case in the current study, has previously been applied to real-time monitoring data (Woods et al., 2020). Exogenous temporal nodes have also been used in prior research (Clasen, Fisher, & Beevers, 2015; Rabinowitz & Fisher, 2020) with similar models as the ones used in the current analyses.

To summarize the GIMME results, we first examined the presence of group-level paths. We next examined the presence and stability of subgroup-level paths. We also sought to describe and characterize these subgroups by examining the path specific to each subgroup. Finally, we describe the individual level paths. To better characterize individual level paths, we also selected a small number of participants and described their individual model and data. For these participants, we provide data visualizations of their raw data to provide readers with a better understanding of how heterogeneity manifests and the connection between the data and the models fit with GIMME. Throughout all analyses, we compare the findings across the adult and adolescent samples. We specifically examine the number of group level paths, the number of subgroups, and the frequency of individual level paths.

For the final aim focused on if different statistical modeling approaches result in different support for or against the IPTS, we used two approaches to contrast with GIMME. We conceptualized this work as a bridge between past research and the current analyses with GIMME. We chose to analyze the same data and variables with alternative statistical approaches to help contextualize the current findings. Both alternative analytic approaches we used had been previously applied to IPTS EMA data (Al-Dajani & Czyz, 2022; Hallensleben et al., 2019; Kleiman et al., 2017; Rath et al., 2019) and therefore serve as helpful comparisons to GIMME.

First, we used multilevel modeling. When referring to multilevel modeling in this manuscript, we are referencing multilevel linear regression. Multilevel modeling pulls all individual effects into a distribution of effects, in contrast to GIMME, where effects are estimated independent of other participants. We specifically ran multilevel models on the same participants used in the GIMME analysis. We focused on the three key theorized relations in the IPTS: burdensomeness to hopeless, isolated to hopeless, and hopeless to urge to kill oneself. Given that the timescale of relationships in the IPTS is unclear (Chu et al., 2017; Millner et al., 2020) and the time lag between observations in the study varied, we focused on contemporaneous associations.

For the multilevel models, we within-person centered all predictors (Hamaker & Muthén, 2020) and used random intercepts and slopes. This model structure allows for both the intercepts and slopes to vary across participants. While more complex versions of multilevel models exist that may better fit the data (Bürkner & Vuorre, 2019; Williams, Mulder, Rouder, & Rast, 2021), to allow for comparison with the regressions used in GIMME we used frequent models with a Gaussian distribution. We fit the models with the *lme4* package (Bates, Mächler, Bolker, & Walker, 2015), computed p-values with the *parameters* package, and plotted marginal effects plots with the *ggeffects* package (Lüdecke, 2018). To visualize heterogeneity within multilevel models, we also plotted individual regressions for each participant within the sample.

The second analytic approach we contrasted with GIMME is multilevel vector autoregression (ml-VAR; Epskamp, Waldorp, Mõttus, & Borsboom, 2018). This analytic approach is often used for modeling multiple relationships between variables and is typically applied within the context of network analysis (Jordan, Winer, & Salem, 2020). ML-VAR models multiple time series at once (i.e., multiple variables from multiple participants) and produces three types of networks: contemporaneous, temporal, and between-person. ML-VAR pools effects across participants so these networks represent the average effect across participants. In the current analysis, we only focus on the contemporaneous and temporal results given that GIMME focuses on within-person relationships. The temporal model represents how each variable predicts itself and other variables over a specific time-lag, typically a lag of one observation. The contemporaneous model shows how variables in network predict each other at the same measurement occasion. All ml-VAR analyses were run separately for the adult and adolescent samples. Analyses were run with *mlVAR* package (Epskamp, Deserno, Bringmann, & Veenman, 2024) and results were plotted with the *qgraph* package (Epskamp et al., 2023). In both samples, we used models with orthogonal random effects rather than correlated random effects based on model fit statistics. In the analyses, only lagged relationships within the same day were examined to restrict the possible duration of lag times.

We briefly summarize the key differences between GIMME, multilevel modeling, and ml-VAR as implemented in the current analyses. First, both multilevel modeling and ml-VAR use pooling to aggregate a mean effect across participants (Haslbeck, Epskamp, & Waldorp, 2023). In GIMME, no multilevel structure is imposed, and no pooling occurs during individual model estimation. Second, GIMME and ml-VAR provide both contemporaneous and lagged/temporal effects for all variables. In this paper, we only modeled contemporaneous relationships in the multilevel modeling analyses. Third, GIMME and ml-VAR model the relationships among all variables in all possible directions. In the multilevel modeling analyses, we only modeled specific theorized relationships in the IPTS. Fourth, the uneven spacing of observations was handled in different ways in GIMME and ml-VAR. For GIMME the time lag duration was included as an exogenous temporal node and for ml-VAR the time-lag was restricted to observations within the same day. Finally, both GIMME and ml-VAR model autoregression, which is how a variable predicts itself over time, and the multilevel modeling approach we used in this study did not. Thus, there are important differences in pooling effects across participants, what relationships between psychological variables are modeled, and how time is modeled in each analysis.

3. Results

3.1. Descriptive statistics

The total number of observations was as follows: adult sample (mean per participant = 227, range = 61 to 495), adolescent sample (mean per participant = 181, range 61 to 470). In the adult sample, the average time in hours between surveys was 7.49 h (SD = 19.85 h) and in the adolescent sample it was 9.18 h (SD = 26.63 h). The distributions of all variables for both samples are provided the supplemental material.

3.2. Group-level paths

When describing paths, we will use language to describe paths such as X to Y. It is important to note that for contemporaneous models these descriptions should be viewed as a description of the statistical relationship of a predictor to an outcome. These associations should not be seen as a causal relationship between variables given the challenges of properly identifying directionality in longitudinal data with weaker autocorrelations (Weigard et al., 2023). Consistent with the IPTS, the hypothesized group-level paths were burdensomeness to hopeless, isolated to hopeless, and hopeless to urge to kill oneself. The presence of group-level paths would support that the IPTS applies to the majority of individuals in the study. In both the adult and adolescent sample, we found zero group level paths.

3.3. Sub-group identification

In the adult sample four subgroups were identified. In the adolescent sample two subgroups were identified. For both samples, the robustness analysis found that very little perturbation (i.e., less than 1% perturbation) results in 20% of community assignments being swapped. This suggests that the subgroups are not stable and therefore should be interpreted with caution (Groen et al., 2022).

In the adult sample, the largest subgroup (n = 45) showed subgroup level paths from burdensomeness to hopeless (contemporaneous), hopeless to feeling isolated (contemporaneous), and hopeless to urge to kill self (contemporaneous). This subgroup showed two out of the three paths hypothesized by the IPTS. The second largest subgroup (n = 33) showed paths from hopeless to burdensomeness (contemporaneous), feeling isolated to feeling hopeless (contemporaneous), and hopeless to urge to kill self (contemporaneous). This subgroup showed two out of the three paths hypothesized by the IPTS. Given that the main difference of these two subgroups are the direction of contemporaneous paths and the uncertainty of direction for contemporaneous paths, caution is warranted when interpreting differences between these two subgroups. The third subgroup (n = 25) did not show associations among IPTS variables and with urge to kill oneself. In this subgroup, there were no significant subgroup level effects. These were individuals that were grouped together by the similarity of their individual models, but no path surpassed the 50% threshold required to be considered subgroup level effect (Wright et al., 2019). The fourth subgroup ($n = 4$) only had one subgroup level path that was a contemporaneous path from hopeless to burdensomeness. The remaining participants in the sample did not belong to any subgroups.

In the adolescent sample, the largest subgroup ($n = 94$) showed subgroup level paths from hopeless to burdensomeness (contemporaneous), feeling isolated to hopeless (contemporaneous), and hopeless to urge to kill self (contemporaneous). This subgroup showed two out of the three paths hypothesized by the IPTS. The second subgroup ($n = 45$) did not show associations among IPTS variables and with urge to kill oneself. In this subgroup, there were no significant subgroup level effects. The remaining participants in the sample did not belong to any subgroups.

3.4. Individual-level paths

There was heterogeneity in the individual-level paths in each study. The frequencies of the paths are shown in Table 3 and Table 4. The percentage of the sample with each path in each sample is shown in Fig. 2.

In the adult sample, the frequencies of the IPTS-implied paths were burdensomeness to hopeless (44.1% contemporaneous; 2.7% lagged), isolated to hopeless (36.0%; 3.6% lagged), and hopeless to urge to kill self (73.0%; 2.7% lagged). The three most common paths were: contemporaneous paths from hopeless to urge to kill self, hopeless to isolated, burdensomeness to hopelessness. Overall, there was a higher prevalence of contemporaneous paths than lagged paths.

In the adolescent sample, the frequencies of the IPTS-implied paths were burdensomeness to hopeless (5.5% contemporaneous; 9.0% lagged), isolated to hopeless (70.3%; 2.1% lagged), and hopeless to urge to kill self (72.4%; 2.8% lagged). The three most common paths were: contemporaneous paths from hopeless to urge to kill self, isolated to hopeless, and hopeless to burdensomeness. Similar to adult sample, there was a higher prevalence of contemporaneous paths than lagged paths.

3.5. Example participants

To provide a deeper understanding of the individual models, we describe two example participants' statistical models and the raw data underlying those models. We highlight one participant that shows a model with paths theorized by the IPTS and one participant that does not show the theorized paths.

Participant 1 is an adolescent participant and completed 186 real-time surveys. The participant had the following significant paths in their model (all paths are positive associations): contemporaneous burdensomeness to isolated, contemporaneous hopeless to burdensomeness, contemporaneous isolated to hopeless, and contemporaneous hopeless to urge to kill self. Thus, the participant has 2 of the three paths proposed by the IPTS. The time series plot in Fig. 3A shows high variability over time in all constructs. The histogram plot in Fig. 3B highlights the zero-inflation in the participant's urge to kill oneself and the high levels of feelings of isolation. The bivariate plots in Fig. 3C–E of the contemporaneous associations hypothesized by the IPTS follow a pattern consistent with the theory. For example, the majority of the times when the participant experienced suicidal urges they were when hopelessness was at a 10 out of 10.

Participant 2 is an adult participant and completed 106 surveys. The participant had the following significant paths in their model: contemporaneous hopeless to feeling isolated (positive association), time in study to feeling isolated (negative association), and time in study to urge to kill self (positive association). The time series plot in Fig. 4A shows moderate variability over time and the temporal trends of decreases in feelings of isolation and increases in suicidal urges. The histogram in Fig. 4B shows the relatively low levels of endorsements of suicidal urges and the high prevalence of 10 out 10 endorsements for all IPTS constructs. The bivariate plots in Fig. 4C–E of the contemporaneous associations hypothesized by the IPTS do not follow a pattern consistent with the theory. There are not linear positive associations among the core constructs. This participant often experiences high levels of burdensomeness, isolation, and hopelessness, and typically experiences lower levels of suicidal urges.

3.6. Multilevel models

In the adult sample, the multilevel model results followed a pattern consistent with the IPTS. There was a significant positive contemporaneousness association between burdensomeness and hopelessness ($B = 0.52$, *95% CI* = 0.46 to 0.58, $p < 0.001$), feeling isolated and hopelessness ($B = 0.44$, *95% CI* = 0.39 to 0.48, $p < 0.001$), and feeling hopelessness and the urge to kill oneself ($B = 0.26$, *95% CI* = 0.21 to 0.31, $p < 0.001$). The bivariate associations are visualized in Fig. 5. Within these overall effects there was heterogeneity as shown in the individual effects in Fig. 5.

In the adolescent sample, the multilevel model results also followed a pattern consistent with the IPTS. There was a significant positive contemporaneous association between burdensomeness and hopelessness ($B = 0.47$, *95% CI* = 0.42 to 0.51, $p < 0.001$), feeling isolated and hopelessness ($B = 0.37$, *95% CI* = 0.33 to 0.41, $p < 0.001$), and feeling hopelessness and the urge to kill oneself ($B = 0.28$, *95% CI* = 0.24 to 0.32, $p < 0.001$). The bivariate associations are visualized in Fig. 6. Within these overall effects there was heterogeneity as shown in the individual effects in Fig. 6.

Table 3
Adult Sample: number of individuals with significant contemporaneous and lagged paths (percent of sample with that path). *Note*: the row is the outcome and the column is predictor in the table.

Contemporaneous Paths	Hopeless	Isolated	Burdensome	Suicidal Urge	Time	Lag
Hopeless	–	40 (36.0)	49 (44.1)	3 (2.7)	4 (3.6)	3 (2.7)
Isolated	51 (45.9)	–	12 (10.8)	6 (5.4)	4 (3.6)	0 (0)
Burdensome	40 (36.0)	11(9.9)	–	3 (2.7)	6 (5.4)	2 (1.8)
Suicidal Urge	81(73.0)	1 (0.9)	4 (3.6)	–	3 (2.7)	2 (1.8)

Lagged Paths	Hopeless	Isolated	Burdensome	Suicidal Urge
Hopeless	111 (100)	4 (3.6)	3 (2.7)	1(0.9)
Isolated	3 (2.7)	111(100)	6 (5.4)	2 (1.8)
Burdensome	5 (4.5)	6 (5.4)	111(100)	3 (2.7)
Suicidal Urge	3 (2.7)	1 (0.9)	3 (2.7)	111(100)

Table 4
Adolescent Sample: number of individuals with significant contemporaneous and lagged paths (percent of sample with that path). *Note*: the row is the outcome and the column is predictor in the table.

Contemporaneous Paths	Hopeless	Isolated	Burdensome	Suicidal Urge	Time	Lag
Hopeless	–	102 (70.3)	8 (5.5)	12 (8.3)	3 (2.1)	3 (2.1)
Isolated	6 (4.1)	–	19 (13.1)	13 (9.0)	7 (4.8)	5 (3.4)
Burdensome	95 (65.5)	17 (11.7)	–	17 (11.7)	8 (5.5)	3 (2.1)
Suicidal Urge	105 (72.4)	6 (4.1)	12 (8.3)	–	7 (4.8)	4 (2.8)

Lagged Paths	Hopeless	Isolated	Burdensome	Suicidal Urge
Hopeless	145 (100)	3 (2.1)	13 (9.0)	1 (0.7)
Isolated	4 (2.8)	145 (100)	1 (0.7)	4 (2.8)
Burdensome	7 (4.8)	1 (0.7)	145 (100)	5 (3.4)
Suicidal Urge	4 (2.8)	5 (3.4)	3 (2.1)	145 (100)

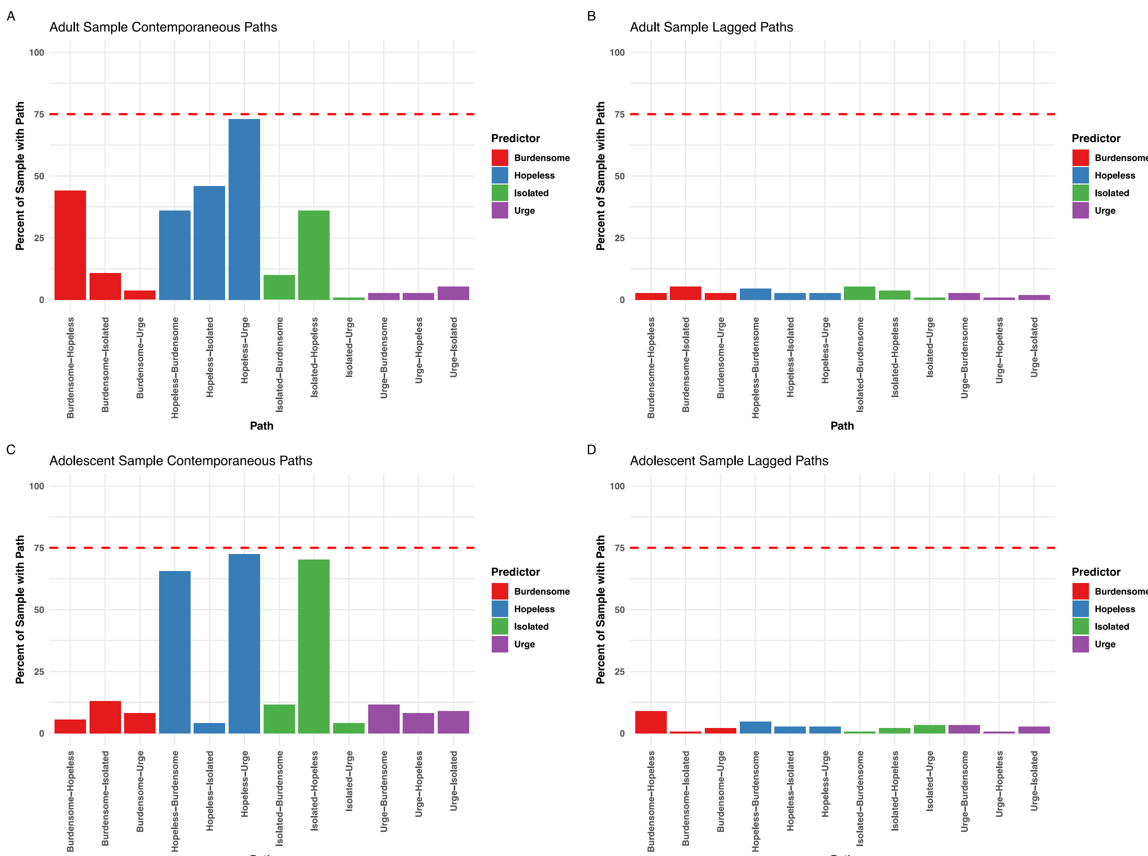

Fig. 2. Percent of Sample with Model Paths. A) Contemporaneous paths in the adult sample; B) Lagged paths in the adult sample; C) Contemporaneous paths in the adolescent sample; D) Lagged paths in the adolescent sample. *Note*: red dashed line at 75% represents the threshold for a group level path. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

3.7. Multi-Level Vector Autoregression Models

The contemporaneous and temporal networks for the adult sample and adolescent sample are shown in Fig. 7. In the contemporaneous networks in both samples, there were significant positive associations among all variables. In the adult sample, the strongest contemporaneous association was between hopelessness and burdensomeness and the weakest was between urge to kill oneself and burdensomeness. In the adolescent sample, the strongest contemporaneous association was between hopelessness and burdensomeness and the weakest was between urge to kill oneself and feelings of isolation. In the temporal networks in both samples, there were significant positive associations among all variables. All variables showed significant positive bidirectional relationships. For example, hopelessness predicted urge to kill oneself and urge to kill oneself predicted hopelessness. The magnitude of all temporal associations was relatively small. The coefficients for all ml-VAR models are provided in the supplemental material.

4. Discussion

We sought to test whether the associations among psychological constructs proposed by well-established theories of suicide are present among most people with suicidal thoughts, or whether there is heterogeneity in these associations across different people. We did this by focusing on the most widely studied current theoretical model of suicide: the IPTS. Across two intensive longitudinal samples, there were three key findings. First, the hypothesis that the associations proposed by this model would hold across most participants was not supported as none of

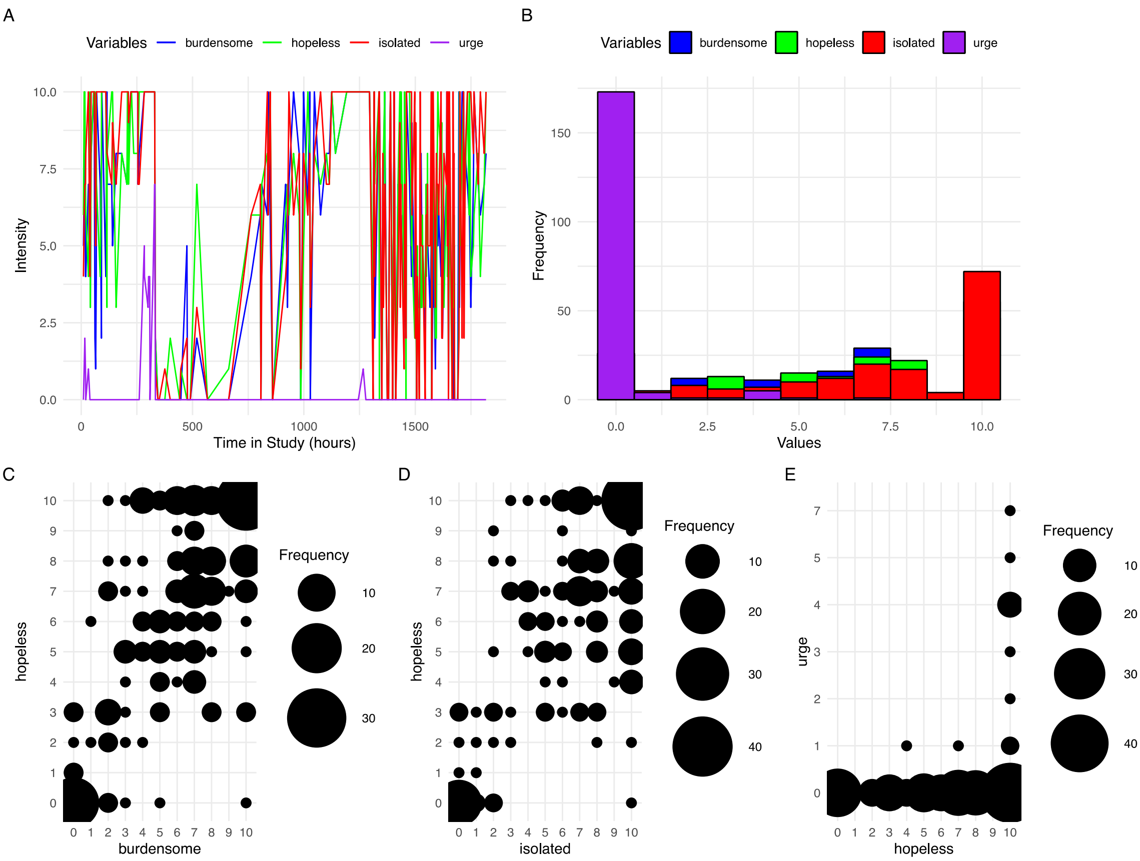

Fig. 3. Example Participant 1. A) Time series of interpersonal constructs and suicidal urges; B) Histogram of interpersonal constructs and suicidal urges; C) Bivariate bubble plot between burdensomeness and hopelessness; D) Bivariate bubble plot between isolation and hopelessness; E) Bivariate bubble plot between hopelessness and suicidal urges.

the theorized paths were significant at the group level. Second, there was evidence of subgroups of individuals who showed similar dynamic models of suicide risk. Third, multilevel models and ml-VAR models, which rely on the average effects across people, showed support for certain relationships proposed by the IPTS whereas GIMME models did not find support. We elaborate on each of these findings below.

No paths appeared for all individuals in the GIMME models. This result does not support the assumption that risk factors apply to all individuals. The current study suggests that existing theories of suicide as they currently stand likely cannot explain the onset of suicidal thinking in all individuals and highlights the challenges of individual level prediction of suicidal thinking (Wang et al., 2023). This study demonstrates how individual level data and approaches such as GIMME allow researchers to quantify the generality of theories. These results suggest that theories such as the IPTS may help to explain suicidal thoughts and behaviors in some individuals, but they are not universal (at least not in the case of the IPTS).

One pathway towards refining theories of suicide is to clarify to whom they apply through subgrouping similar individuals. In this study, we identified subgroups of individuals based on dynamic within-person patterns. These subgroups, however, were not stable. Prior research has identified subgroups of suicide risk in cross-sectional data (Ginley & Bagge, 2017) and subgroups of patterns of suicidal thinking based on within-person data (Kleiman et al., 2018). Recently theoretical work has proposed the importance of multiple suicidal subtypes (Bernanke, Stanley, & Oquendo, 2017). Our findings advance this area of research by combining intensive within-person data on suicidal thinking risk with subgrouping algorithms. In both the adult and adolescent sample, there was a subgroup of individuals with contemporaneous patterns consistent with the IPTS. These findings suggest that theories, such as IPTS, appear unlikely to currently represent the cause of suicide for all people, but may be relevant in the pathway to suicide for some. The lack of stability of the subgroups in the current study and other studies applying similar methods (Ellison et al., 2020; Groen et al., 2022) or lack of subgroup level effects (Kaurin et al., 2021), highlights the difficulty of identifying meaningful subgroups in the face of immense heterogeneity. Nevertheless, leveraging new dynamic data and algorithms to identify to whom theories apply is a promising way of improving the clinical usefulness of current theories.

The final aim was to compare if traditional statistical approaches, which rely on the average effects across people, result in different support for or against the IPTS than GIMME which relies on individual level effects. GIMME found low support for the IPTS with no group level contemporaneous or lagged paths. We found that when the same data were examined with multilevel modeling, there were significant contemporaneous associations for all the theorized relationships within the IPTS. With multilevel vector autoregression, there were also significant contemporaneous associations among all IPTS variables. There were also significant temporal associations of theorized relationships, but there were also paths in the other (non-theorized direction) and all temporal effects were relatively small. Thus, these three different approaches all provide different levels of support for the IPTS. The notion that the same data when examined with idiographic methods (i.e., GIMME) versus multilevel methods (i.e., ml-VAR) could provide different results may seem surprising but is consistent with past work in other areas (Sahdra et al., 2024). Consistent with past general meta-scientific work (Bastiaansen et al., 2020; Silberzahn et al., 2018), these findings emphasize the impact of different analytic decisions in

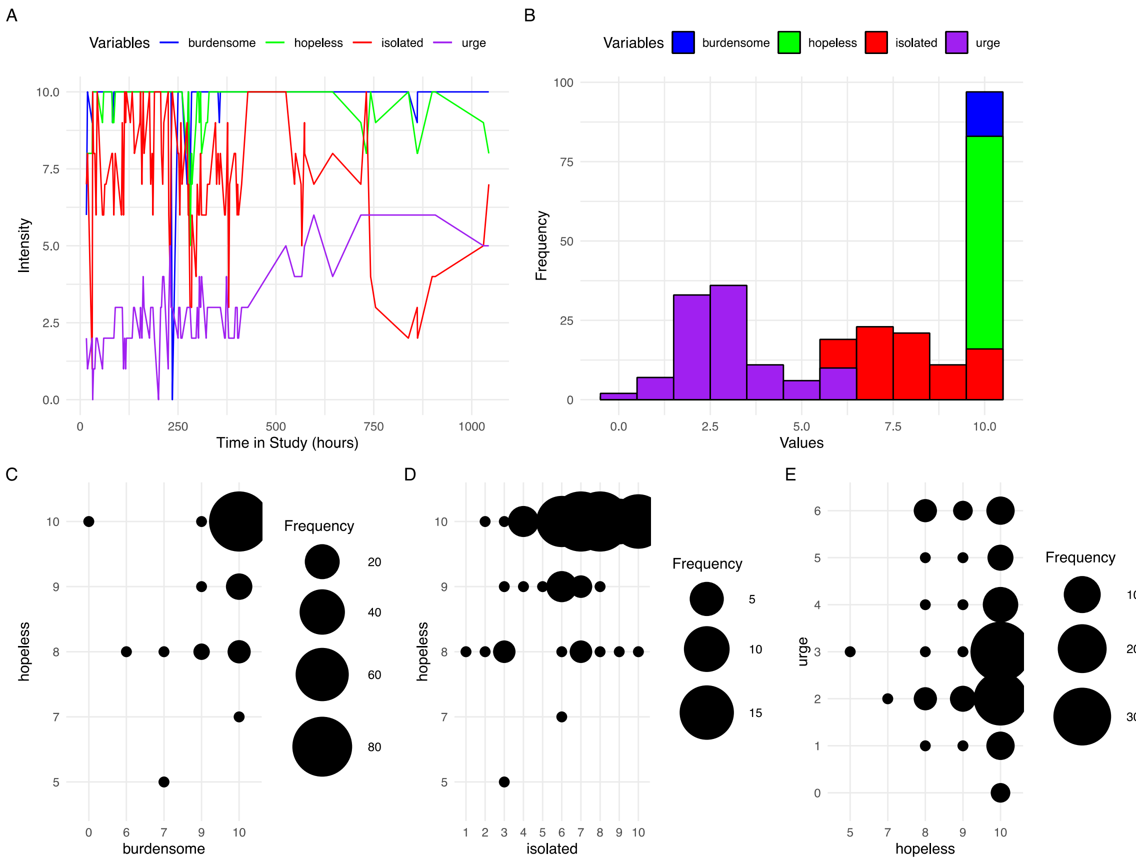

Fig. 4. Example Participant 2. A) Time series of interpersonal constructs and suicidal urges; B) Histogram of interpersonal constructs and suicidal urges; C) Bivariate bubble plot between burdensomeness and hopelessness; D) Bivariate bubble plot between isolation and hopelessness; E) Bivariate bubble plot between hopelessness and suicidal urges.

supporting hypotheses. As noted in the methods section, all of these analytic approaches handled heterogeneity and time in different ways and these differences seem to have significant implications. This finding also demonstrates possible discrepancies between a general effect in a population and an effect at the individual level (Fisher et al., 2018). This discrepancy is important because psychological theories are currently vague about intraindividual effects and interindividual effects (Borsboom & Haslbeck (2024)). An overall implication of this finding is the need for theories to be more specific (Millner et al., 2020; Robinaugh, Haslbeck, Ryan, Fried, & Waldorp, 2021) reducing the gap between theories and the statistical models used to test them.

While the present study has several strengths such as ecological validity, within-person data, and use of multiple samples, there are several key limitations. First, participants across both samples were primarily White, cisgender individuals. This limits the generality of the findings, especially given that specific theories may be more relevant among certain demographic groups (Opara et al., 2020). Second, the study only focused on suicidal thoughts, but the IPTS also makes predictions for suicidal behaviors (Van Orden et al., 2010). Capturing suicidal behavior in real-time monitoring is challenging (Coppersmith, Wang, et al., 2023; Rogers, 2023), given the low-base rate of the behavior. Still the current study is only able to speak to the generality and validity of the IPTS on suicidal thoughts. Third, the study did not test interactions between psychological constructs in the IPTS. Tests of the IPTS often focus on interactions among core constructs, such as perceived burdensomeness and thwarted belongingness (Chu et al., 2017). Testing interactions in GIMME is possible (Arizmendi et al., 2020), but requires adequate statistical power rarely available with idiographic models. Fourth, we did not use multiple-solutions GIMME (Beltz & Molenaar, 2016; Weigard et al., 2023), which can be useful for improving inference about the directionality of paths. Multiple-solutions GIMME, however, is not compatible with S-GIMME (Weigard et al., 2023). Fifth, the current study used single items to assess suicidal thoughts and IPTS constructs. This could have introduced measurement error (Dejonckheere et al., 2022) and impacted the ability to uncover associations. Finally, the design of the real-time monitoring studies may have failed to capture lagged effects. There were far more significant contemporaneous than lagged effects across studies, which is consistent with other research applying GIMME (Woods et al., 2020). This may be because effects may show up as contemporaneous when assessments occur at a slower timescale than the change in the measured phenomenon (Granger, 1969; Woods et al., 2020). Future work with higher density sampling (e.g., minutes apart) would be better able to disentangle contemporaneous versus lagged effects (Coppersmith, Ryan, et al., 2023).

This project has important implications for understanding suicide. First, the current finding is based on a data-driven approach that aims to reveal meaningful relationships between critical constructs and how these relationships vary within subgroups. Yet such approaches may be picking up on "surface-level heterogeneity" (i.e., noise) (Wright & Woods, 2020). If this is the case, integrating data-driven approaches and theoretical frameworks may advance understanding more than data-driven approaches alone. Second, the heterogeneity identified in this study and other work (Kaurin et al., 2021) poses a challenge for theories that posit a common final pathway for all suicidal behaviors (Millner, Lee, & Nock, 2017)

This work also has important implications for suicide prevention. Future interventions should accommodate the heterogeneity in risk

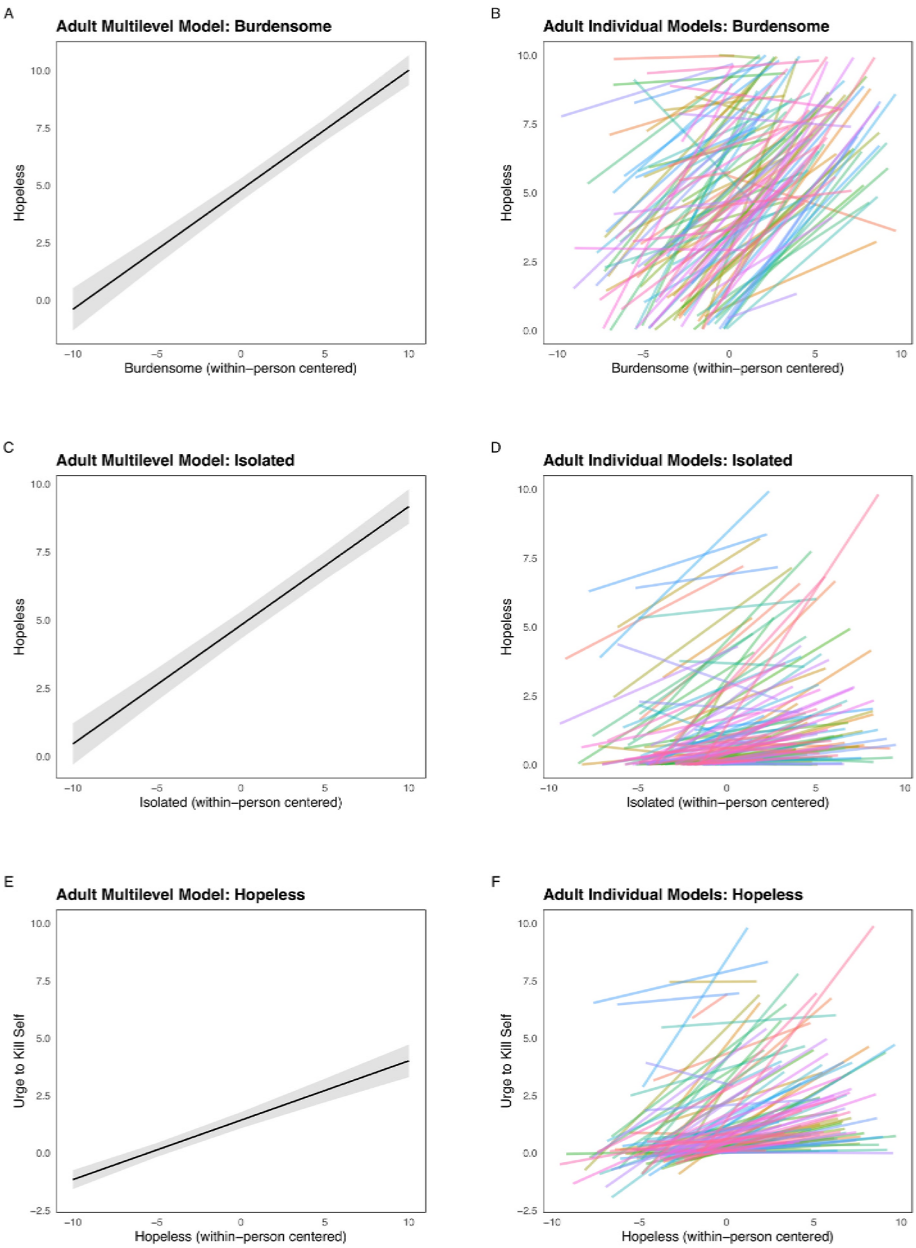

Fig. 5. Multilevel model results adult sample.

captured in the present study. This heterogeneity in risk may be driving heterogeneity in treatment effects (Kessler, 2019) due to the homogenous nature of existing suicide prevention efforts. There are two future ways suicide prevention efforts can account for this heterogeneity. First, treatment decision algorithms can be applied to allocate specific existing interventions based on predictions of what intervention is most likely to be helpful for individual patients (DeRubeis, 2019; Kessler, Chalker, Luedtke, Sadikova, & Jobes, 2020). Second, it may be that interventions work for certain people only at certain times. Thus, just-in-time adaptive interventions (JITAIs) are designed to provide the right type of support at the right time by adapting to changes in internal states and external contexts (Nahum-Shani et al., 2018). JITAIs may match the dynamic, heterogeneous, and interactive nature of suicidal thoughts and behaviors (Coppersmith et al., 2022).

The current study provides previously unavailable information about the heterogeneity of existing suicide risk factors. We hope this work serves as a catalyst for advancing suicide theory and prevention.

Funding statement

This work was funded by the National Institute of Mental Health Grant U01-MH116928. Daniel D.L. Coppersmith is supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1745303 and the National Institute of Mental Health under Grant F31MH130055. Shirley B. Wang is supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1745303 and the National Institute of Mental Health under Grant F31MH125495. Rebecca G. Fortgang is supported by the National Institute of Mental Health under Grant K23MH132766. Kate H. Bentley is supported by the National Institute of Mental Health under Grant K23MH120436. This content is solely the responsibility of the authors and does not necessarily represent the official views of the National Science Foundation or the National Institutes of Health.

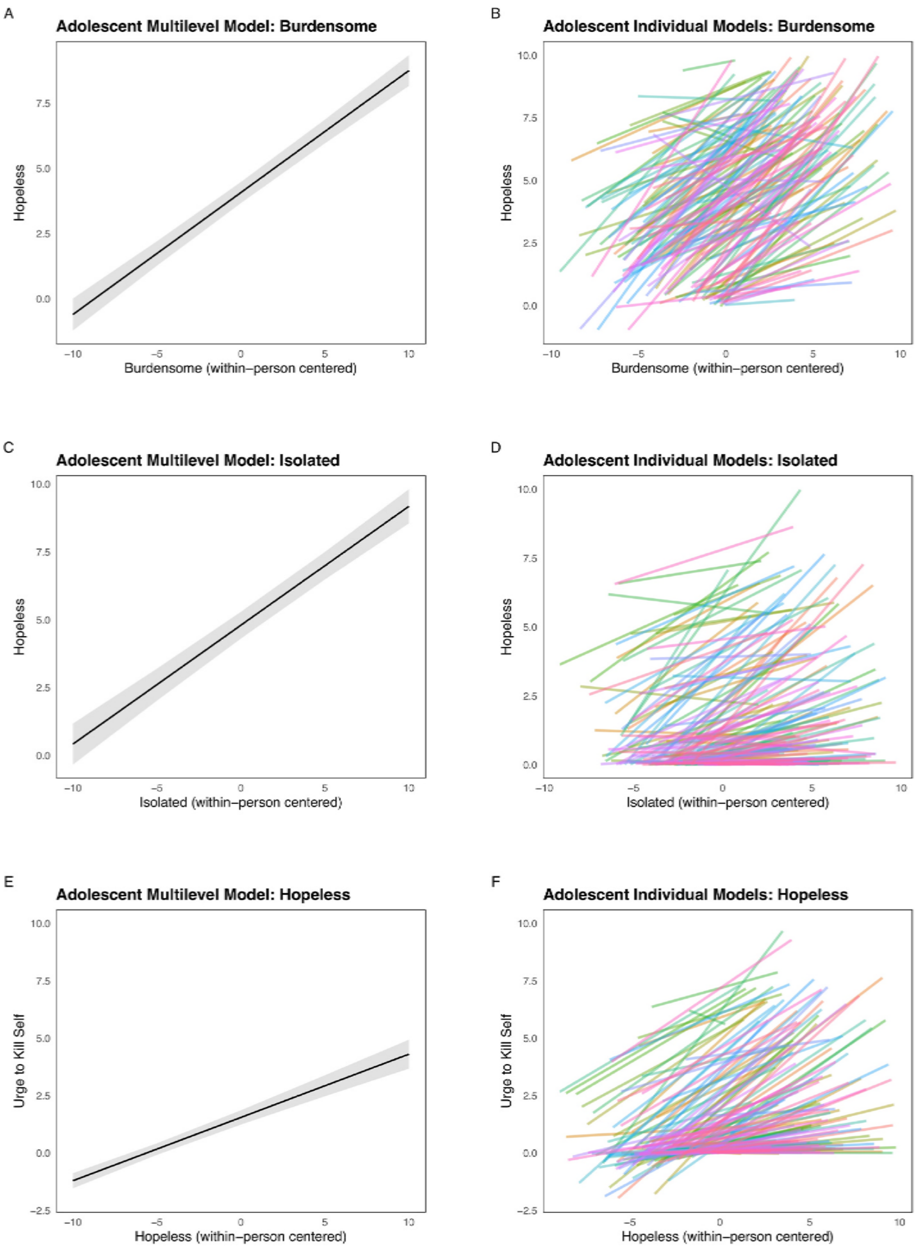

Fig. 6. Multilevel model results adolescent sample.

CRediT authorship contribution statement

Daniel D.L. Coppersmith: Conceptualization, Formal analysis, Writing – original draft. **Evan M. Kleiman:** Conceptualization, Investigation, Writing – original draft. **Alexander J. Millner:** Conceptualization, Investigation, Writing – review & editing. **Shirley B. Wang:** Writing – review & editing. **Cara Arizmendi:** Formal analysis, Software, Writing – review & editing. **Kate H. Bentley:** Investigation, Writing – review & editing. **Dylan DeMarco:** Data curation, Writing – review & editing. **Rebecca G. Fortgang:** Investigation, Writing – review & editing. **Kelly L. Zuromski:** Investigation, Writing – review & editing. **Joseph S. Maimone:** Writing – review & editing. **Adam Haim:** Writing – review & editing. **Jukka-Pekka Onnela:** Investigation, Writing – review & editing. **Suzanne A. Bird:** Investigation, Resources, Writing – review & editing. **Jordan W. Smoller:** Investigation, Resources, Writing – review & editing. **Patrick Mair:** Formal analysis, Writing – review & editing. **Matthew K. Nock:** Conceptualization, Funding acquisition, Investigation, Project administration, Resources, Supervision, Writing – review & editing.

Declaration of competing interest

Dr. Nock receives publication royalties from Macmillan, Pearson, and UpToDate. He has been a paid consultant in the past three years for Microsoft Corporation, the Veterans Health Administration, and COMPASS Pathways, and for legal cases regarding a death by suicide. He has stock options in Cerebral Inc. He is an unpaid scientific advisor for Empatica, Koko, and TalkLife. Dr. Kleiman has been a paid consultant in the past three years for Boehringer Ingelheim Pharmaceuticals. Dr. Smoller is a member of the Scientific Advisory Board of Sensorium Therapeutics (with equity), and has received grant support from Biogen, Inc. He is PI of a collaborative study of the genetics of depression and bipolar disorder sponsored by 23andMe for which 23andMe provides analysis time as in-kind support but no payments. Dr. Onnela is a cofounder and board member of a commercial entity that operates in digital phenotyping.

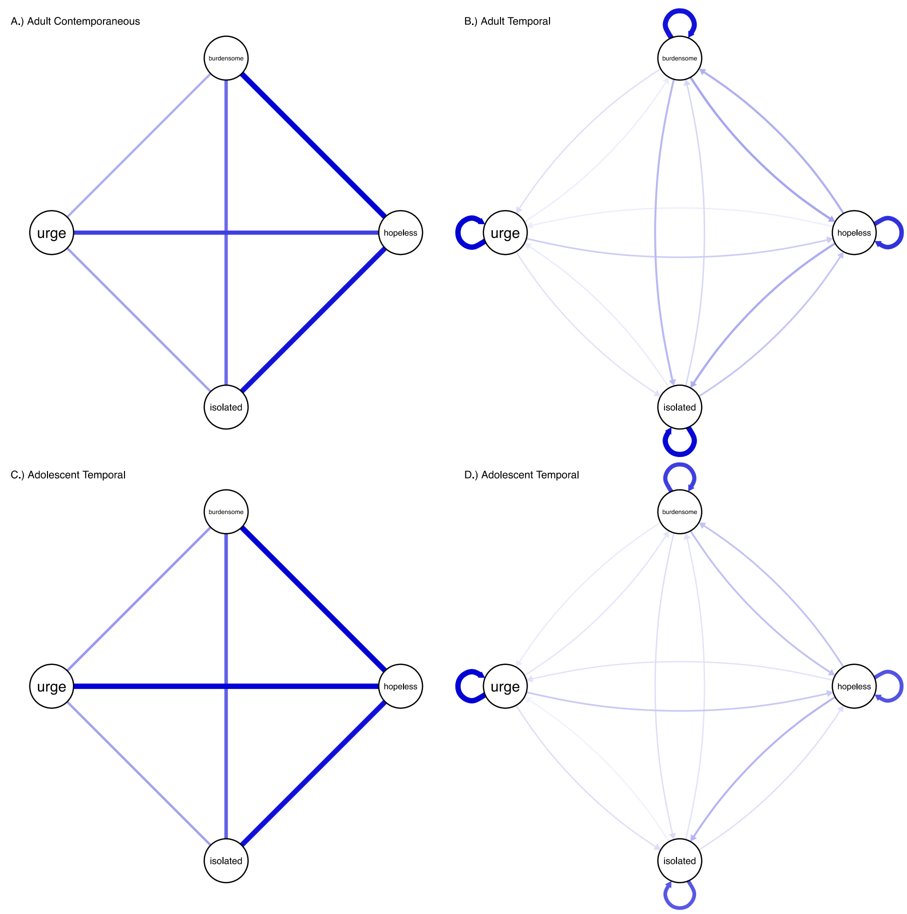

Fig. 7. Multi-Level Vector Autoregression Models. A) Contemporaneous Network in the Adult Sample; B) Temporal Network in the Adult Sample; C) Contemporaneous Network in the Adolescent Sample; D) Temporal Network in the Adolescent Sample. *Note*: Blue edges indicate a positive association, and the thickness of the edge corresponds to the strength of the associations. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Data availability

Access to anonymized data will be available through the National Institute of Mental Health Data Archive at the completion of data collection for the larger project of which this study is a component.

References

Al-Dajani, N., & Czyz, E. K. (2022). Suicidal desire in adolescents: An examination of the interpersonal psychological theory using daily diaries. *Journal of Clinical Child and Adolescent Psychology, 53*, 1–15. https://doi.org/10.1080/15374416.2022.2051525. *The Official Journal for the Society of Clinical Child and Adolescent Psychology, American Psychological Association, Division*.

Arizmendi, C., Gates, K., Fredrickson, B., & Wright, A. (2020). Specifying exogeneity and bilinear effects in data-driven model searches. *Behavior Research Methods*. https://doi.org/10.3758/s13428-020-01469-2

Borsboom, D., & Haslbeck, J. (2024). Integrating intra- and interindividual phenomena in psychological theories. https://doi.org/10.31234/osf.io/az9qk.

Bürkner, P.-C., & Vuorre, M. (2019). Ordinal regression models in psychology: A tutorial. *Advances in Methods and Practices in Psychological Science, 2*(1), 77–101. https://doi.org/10.1177/2515245918823199

Barlow, D. H., & Nock, M. K. (2009). Why can't we Be more idiographic in our research? *Perspectives on Psychological Science, 4*(1), 19–21. https://doi.org/10.1111/j.1745-6924.2009.01088.x

Bastiaansen, J. A., Kunkels, Y. K., Blaauw, F. J., Boker, S. M., Ceulemans, E., Chen, M., et al. (2020). Time to get personal? The impact of researchers choices on the selection of treatment targets using the experience sampling methodology. *Journal of Psychosomatic Research, 137*, Article 110211. https://doi.org/10.1016/j.jpsychores.2020.110211

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

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., & Molenaar, P. C. M. (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

Bentley, K. H., Coppersmith, D. L., Kleiman, E. M., Nook, E. C., Mair, P., Millner, A. J., et al. (2021). Do patterns and types of negative affect during hospitalization predict short-term post-discharge suicidal thoughts and behaviors? *Affective Science, 2*(4), 484–494. https://doi.org/10.1007/s42761-021-00058-6

Bernanke, J. A., Stanley, B. H., & Oquendo, M. A. (2017). Toward fine-grained phenotyping of suicidal behavior: The role of suicidal subtypes. *Molecular Psychiatry, 22*(8), 1080–1081. https://doi.org/10.1038/mp.2017.123

Christensen, H., Batterham, P. J., Soubelet, A., & Mackinnon, A. J. (2013). A test of the Interpersonal Theory of Suicide in a large community-based cohort. *Journal of Affective Disorders, 144*(3), 225–234. https://doi.org/10.1016/j.jad.2012.07.002

Chu, C., Buchman-Schmitt, J. M., Stanley, I. H., Hom, M. A., Tucker, R. P., Hagan, C. R., et al. (2017). The interpersonal theory of suicide: A systematic review and meta-analysis of a decade of cross-national research. *Psychological Bulletin, 143*(12), 1313–1345. https://doi.org/10.1037/bul0000123

Chu, C., Zuromski, K. L., Bernecker, S. L., Gutierrez, P. M., Joiner, T. E., Liu, H., et al. (2020). A test of the interpersonal theory of suicide in a large, representative, retrospective and prospective study: Results from the Army Study to Assess Risk and Resilience in Servicemembers (Army STARRS). *Behaviour Research and Therapy, 132*, Article 103688. https://doi.org/10.1016/j.brat.2020.103688

Clasen, P. C., Fisher, A. J., & Beevers, C. G. (2015). Mood-reactive self-esteem and depression vulnerability: Person-specific symptom dynamics via smart phone assessment. *PLoS One, 10*(7), Article e0129774. https://doi.org/10.1371/journal.pone.0129774

Coppersmith, D. D. L., Dempsey, W., Kleiman, E. M., Bentley, K. H., Murphy, S. A., & Nock, M. K. (2022). Just-in-Time adaptive interventions for suicide prevention: Promise, challenges, and future directions. *Psychiatry, 85*(4), 317–333. https://doi.org/10.1080/00332747.2022.2092828

Coppersmith, D. D. L., Millgram, Y., Kleiman, E. M., Fortgang, R. G., Millner, A. J., Frumkin, M. R., et al. (2023). Suicidal thinking as affect regulation. *Journal of*

Psychopathology and Clinical Science, 132(4), 385–395. https://doi.org/10.1037/abn0000828

Coppersmith, D. D. L., Ryan, O., Fortgang, R. G., Millner, A. J., Kleiman, E. M., & Nock, M. K. (2023). Mapping the timescale of suicidal thinking. *Proceedings of the National Academy of Sciences, 120*(17), Article e2215434120. https://doi.org/10.1073/pnas.2215434120

Coppersmith, D. D. L., Wang, S. B., Kleiman, E. M., Maimone, J. S., Fedor, S., Bentley, K. H., et al. (2023). Real-time digital monitoring of a suicide attempt by a hospital patient. *General Hospital Psychiatry, 80*, 35–39. https://doi.org/10.1016/j.genhosppsych.2022.12.005

Czyz, E. K., Horwitz, A. G., Arango, A., & King, C. A. (2019). Short-term change and prediction of suicidal ideation among adolescents: A daily diary study following psychiatric hospitalization. *Journal of Child Psychology and Psychiatry, 60*(7), 732–741. https://doi.org/10.1111/jcpp.12974

Dejonckheere, E., Demeyer, F., Geusens, B., Piot, M., Tuerlinckx, F., Verdonck, S., et al. (2022). Assessing the reliability of single-item momentary affective measurements in experience sampling. *Psychological Assessment, 34*(12), 1138–1154. https://doi.org/10.1037/pas0001178

DeRubeis, R. J. (2019). The history, current status, and possible future of precision mental health. *Behaviour Research and Therapy, 123*, Article 103506. https://doi.org/10.1016/j.brat.2019.103506

Dokoupil, T. (2013). In *The suicide epidemic*. Newsweek https://www.newsweek.com/2013/05/22/why-suicide-has-become-epidemic-and-what-we-can-do-help-237434.html.

Ellison, W. D., Levy, K. N., Newman, M. G., Pincus, A. L., Wilson, S. J., & Molenaar, P. C. M. (2020). Dynamics among borderline personality and anxiety features in psychotherapy outpatients: An exploration of nomothetic and idiographic patterns. *Personality Disorders, 11*(2), 131–140. https://doi.org/10.1037/per0000363

Epskamp, S., Costantini, G., Haslbeck, J., Isvoranu, A., Cramer, A. O. J., Waldorp, L. J., et al. (2023). In *qgraph: Graph plotting methods, psychometric data visualization and graphical model estimation (1.9.8)* [Computer software] https://cran.r-project.org/web/packages/qgraph/index.html.

Epskamp, S., Deserno, M. K., Bringmann, L. F., & Veenman, M. (2024). mlVAR: Multi-Level vector autoregression (0.5.2) [Computer software] https://cran.r-project.org/web/packages/mlVAR/index.html.

Epskamp, S., Waldorp, L. J., Mõttus, R., & Borsboom, D. (2018). The Gaussian graphical model in cross-sectional and time-series data. *Multivariate Behavioral Research, 53*(4), 453–480. https://doi.org/10.1080/00273171.2018.1454823

Fisher, A. J. (2015). Toward a dynamic model of psychological assessment: Implications for personalized care. *Journal of Consulting and Clinical Psychology, 83*(4), 825–836. https://doi.org/10.1037/ccp0000026

Fisher, A. J., Medaglia, J. D., & Jeronimus, B. F. (2018). Lack of group-to-individual generalizability is a threat to human subjects research. *Proceedings of the National Academy of Sciences, 115*(27), E6106–E6115. https://doi.org/10.1073/pnas.1711978115

Fisher, A. J., Reeves, J. W., Lawyer, G., Medaglia, J. D., & Rubel, J. A. (2017). Exploring the idiographic dynamics of mood and anxiety via network analysis. *Journal of Abnormal Psychology, 126*(8), 1044–1056. https://doi.org/10.1037/abn0000311

Fox, K. R., Harris, J. A., Wang, S. B., Millner, A. J., Deming, C. A., & Nock, M. K. (2020). Self-injurious thoughts and behaviors interview-revised: Development, reliability, and validity. *Psychological Assessment*. https://doi.org/10.1037/pas0000819

Gates, K. M., Fisher, Z. F., Arizmendi, C., Henry, T. R., Duffy, K. A., & Mucha, P. J. (2019). Assessing the robustness of cluster solutions obtained from sparse count matrices. *Psychological Methods, 24*(6), 675–689. https://doi.org/10.1037/met0000204

Gates, K. M., Lane, S. T., Varangis, E., Giovanello, K., & Guskiewicz, K. (2017). Unsupervised classification during time-series model building. *Multivariate Behavioral Research, 52*(2), 129–148. https://doi.org/10.1080/00273171.2016.1256187

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

Gee, B. L., Han, J., Benassi, H., & Batterham, P. J. (2020). Suicidal thoughts, suicidal behaviours and self-harm in daily life: A systematic review of ecological momentary assessment studies. *Digital Health, 6*, Article 2055207620963958. https://doi.org/10.1177/2055207620963958

Gelman, A., Hill, J., & Yajima, M. (2012). Why we (usually) don't have to worry about multiple comparisons. *Journal of Research on Educational Effectiveness, 5*(2), 189–211. https://doi.org/10.1080/19345747.2011.618213

Ginley, M. K., & Bagge, C. L. (2017). Psychiatric heterogeneity of recent suicide attempters: A latent class analysis. *Psychiatry Research, 251*, 1–7. https://doi.org/10.1016/j.psychres.2017.02.004

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica, 37*(3), 424–438. https://doi.org/10.2307/1912791. JSTOR.

Groen, R. N., Arizmendi, C., Wichers, M., Schreuder, M. J., Gates, K. M., Hartman, C. A., et al. (2022). Shared and individual-specific daily stress-reactivity in a cross-diagnostic at-risk sample. *Journal of Psychopathology and Clinical Science, 131*(3), 221–234. https://doi.org/10.1037/abn0000745

Hallensleben, N., Glaesmer, H., Forkmann, T., Rath, D., Strauss, M., Kersting, A., et al. (2019). Predicting suicidal ideation by interpersonal variables, hopelessness and depression in real-time. An ecological momentary assessment study in psychiatric inpatients with depression. *European Psychiatry, 56*(1), 43–50. https://doi.org/10.1016/j.eurpsy.2018.11.003

Hamaker, E. L. (2012). Why researchers should think "within-person": A paradigmatic rationale. In *Handbook of research methods for studying daily life* (pp. 43–61) (Guilford Press).

Hamaker, E. L., & Muthén, B. (2020). The fixed versus random effects debate and how it relates to centering in multilevel modeling. *Psychological Methods, 25*(3), 365–379. https://doi.org/10.1037/met0000239

Hamaker, E. L., & Wichers, M. (2017). No time like the present: Discovering the hidden dynamics in intensive longitudinal data. *Current Directions in Psychological Science, 26*(1), 10–15. https://doi.org/10.1177/0963721416666518

Haslbeck, J., Epskamp, S., & Waldorp, L. (2023). *Testing for group differences in multilevel vector autoregressive models*. https://doi.org/10.31234/osf.io/dhp8s. OSF.

Hausmann-Stabile, C., Glenn, C. R., & Kandlur, R. (2021). Theories of suicidal thoughts and behaviors: What exists and what is needed to advance youth suicide research. In R. Miranda, & E. L. Jeglic (Eds.), *Handbook of youth suicide prevention: Integrating research into practice* (pp. 9–29). https://doi.org/10.1007/978-3-030-82465-5_2. Springer International Publishing.

Hjelmeland, H., & Loa Knizek, B. (2020). The emperor's new clothes? A critical look at the interpersonal theory of suicide. *Death Studies, 44*(3), 168–178. https://doi.org/10.1080/07481187.2018.1527796

Joiner, T. (2007). *Why people die by suicide*. Harvard University Press.

Joiner, T. E., Jeon, M. E., Lieberman, A., Janakiraman, R., Duffy, M. E., Gai, A. R., et al. (2021). On prediction, refutation, and explanatory reach: A consideration of the interpersonal theory of suicidal behavior. *Preventive Medicine, 152*, Article 106453. https://doi.org/10.1016/j.ypmed.2021.106453

Jordan, D. G., Winer, E. S., & Salem, T. (2020). The current status of temporal network analysis for clinical science: Considerations as the paradigm shifts? *Journal of Clinical Psychology, 76*(9), 1591–1612. https://doi.org/10.1002/jclp.22957

Kaurin, A., Dombrovski, A. Y., Hallquist, M. N., & Wright, A. G. C. (2021). Integrating a functional view on suicide risk into idiographic statistical models. *Behaviour Research and Therapy, 104012*. https://doi.org/10.1016/j.brat.2021.104012

Kessler, R. C. (2019). Clinical epidemiological research on suicide-related behaviors—where we are and where we need to go. *JAMA Psychiatry, 76*(8), 777–778. https://doi.org/10.1001/jamapsychiatry.2019.1238

Kessler, R. C., Chalker, S. A., Luedtke, A. R., Sadikova, E., & Jobes, D. A. (2020). A preliminary precision treatment rule for remission of suicide ideation. *Suicide and Life-Threatening Behavior, 50*(2), 558–572. https://doi.org/10.1111/sltb.12609

Kim, J., Zhu, W., Chang, L., Bentler, P. M., & Ernst, T. (2007). Unified structural equation modeling approach for the analysis of multisubject, multivariate functional MRI data. *Human Brain Mapping, 28*(2), 85–93. https://doi.org/10.1002/hbm.20259

Kleiman, E. M., Glenn, C. R., & Liu, R. T. (2023). The use of advanced technology and statistical methods to predict and prevent suicide. *Nature Reviews Psychology, 2*(6). https://doi.org/10.1038/s44159-023-00175-y. Article 6.

Kleiman, E. M., Turner, B. J., Fedor, S., Beale, E. E., Huffman, J. C., & Nock, M. K. (2017). Examination of real-time fluctuations in suicidal ideation and its risk factors: Results from two ecological momentary assessment studies. *Journal of Abnormal Psychology, 126*(6), 726–738. https://doi.org/10.1037/abn0000273

Kleiman, E. M., Turner, B. J., Fedor, S., Beale, E. E., Picard, R. W., Huffman, J. C., et al. (2018). Digital phenotyping of suicidal thoughts. *Depression and Anxiety, 35*(7), 601–608. https://doi.org/10.1002/da.22730

Lüdecke, D. (2018). ggeffects: Tidy data frames of marginal effects from regression models. *Journal of Open Source Software, 3*(26), 772. https://doi.org/10.21105/joss.00772

Lane, S., Gates, K., Fisher, Z., Arizmendi, C., Molenaar, P., Hallquist, M., et al. (2021). *gimme: Group iterative multiple model estimation* (0.7-6) [Computer software] https://CRAN.R-project.org/package=gimme.

Lane, S., 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), 54–69. https://doi.org/10.1037/met0000192

Liu, S. (2017). Person-specific versus multilevel autoregressive models: Accuracy in parameter estimates at the population and individual levels. *British Journal of Mathematical and Statistical Psychology, 70*(3), 480–498. https://doi.org/10.1111/bmsp.12096

Miller, A. B., & Prinstein, M. J. (2019). Adolescent suicide as a failure of acute stress-response systems. *Annual Review of Clinical Psychology, 15*, 425–450. https://doi.org/10.1146/annurev-clinpsy-050718-095625

Millner, A. J., Lee, M. D., & Nock, M. K. (2017). Describing and measuring the pathway to suicide attempts: A preliminary study. *Suicide and Life-Threatening Behavior, 47*(3), 353–369. https://doi.org/10.1111/sltb.12284

Millner, A. J., Robinaugh, D. J., & Nock, M. K. (2020). Advancing the understanding of suicide: The need for formal theory and rigorous descriptive research. *Trends in Cognitive Sciences*. https://doi.org/10.1016/j.tics.2020.06.007

Minois, G. (2001). *History of suicide: Voluntary Death in western culture (L. G. Cochrane, trans.)*. Johns Hopkins University Press.

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 and Perspectives, 2*(4), 201–218. https://doi.org/10.1207/s15366359mea0204_1

Nahum-Shani, I., Smith, S. N., Spring, B. J., Collins, L. M., Witkiewitz, K., Tewari, A., et al. (2018). Just-in-Time adaptive interventions (JITAIs) in mobile health: Key components and design principles for ongoing health behavior support. *Annals of Behavioral Medicine: A Publication of the Society of Behavioral Medicine, 52*(6), 446–462. https://doi.org/10.1007/s12160-016-9830-8

Nock, M. K., Borges, G., Bromet, E. J., Cha, C. B., Kessler, R. C., & Lee, S. (2008). Suicide and suicidal behavior. *Epidemiologic Reviews, 30*(1), 133–154. https://doi.org/10.1093/epirev/mxn002

Nock, M. K., Holmberg, E. B., Photos, V. I., & Michel, B. D. (2007). Self-injurious thoughts and behaviors Interview: Development, reliability, and validity in an adolescent sample. *Psychological Assessment, 19*(3), 309–317. https://doi.org/10.1037/1040-3590.19.3.309

Opara, I., Assan, M. A., Pierre, K., Gunn, J. F., Metzger, I., Hamilton, J., et al. (2020). Suicide among black children: An integrated model of the interpersonal-psychological theory of suicide and intersectionality theory for researchers and clinicians. *Journal of Black Studies, 51*(6), 611–631. https://doi.org/10.1177/0021934720935641

Rabinowitz, A. R., & Fisher, A. J. (2020). Person-specific methods for characterizing the course and temporal dynamics of concussion symptomatology: A pilot study. *Scientific Reports, 10*(1), 1–9. https://doi.org/10.1038/s41598-019-57220-1

Rath, D., de Beurs, D., Hallensleben, N., Spangenberg, L., Glaesmer, H., & Forkmann, T. (2019). Modelling suicide ideation from beep to beep: Application of network analysis to ecological momentary assessment data. *Internet Interventions, 18*, Article 100292. https://doi.org/10.1016/j.invent.2019.100292

Robinaugh, D. J., Haslbeck, J. M. B., Ryan, O., Fried, E. I., & Waldorp, L. J. (2021). Invisible hands and fine calipers: A call to use formal theory as a toolkit for theory construction. *Perspectives on Psychological Science, 1745691620974697*. https://doi.org/10.1177/1745691620974697

Rogers, M. L. (2023). A real-time evaluation of within-person and between-person risk for suicidal behaviors. *International Journal of Cognitive Therapy*. https://doi.org/10.1007/s41811-023-00174-0

Sahdra, B. K., Ciarrochi, J., Klimczak, K. S., Krafft, J., Hayes, S. C., & Levin, M. (2024). Testing the applicability of idionomic statistics in longitudinal studies: The example of 'doing what matters. *Journal of Contextual Behavioral Science, 32*, Article 100728. https://doi.org/10.1016/j.jcbs.2024.100728

Selby, E., Jr, T. E. J., & Ribeiro, J. D. (2014). Comprehensive theories of suicidal behaviors. https://doi.org/10.1093/oxfordhb/9780195388565.013.0016.

Sewall, C. J. R., & Wright, A. G. C. (2021). Personalizing suicidology. *Crisis, 42*(6), 405–410. https://doi.org/10.1027/0227-5910/a000834

Silberzahn, R., Uhlmann, E. L., Martin, D. P., Anselmi, P., Aust, F., Awtrey, E., et al. (2018). Many analysts, one dataset: Making transparent how variations in analytical choices affect results. *Advances in Methods and Practices in Psychological Science, 1*(3), 337–356. https://doi.org/10.1177/2515245917747646

Stone, A. A., Schneider, S., & Smyth, J. M. (2023). Evaluation of pressing issues in ecological momentary assessment. *Annual Review of Clinical Psychology, 19*, 107–131. https://doi.org/10.1146/annurev-clinpsy-080921-083128

Van Orden, K. A., Witte, T. K., Cukrowicz, K. C., Braithwaite, S., Selby, E. A., & Joiner, T. E. (2010). The interpersonal theory of suicide. *Psychological Review, 117*(2), 575–600. https://doi.org/10.1037/a0018697

Wang, S. B., Coppersmith, D. D. L., Kleiman, E. M., Bentley, K. H., Millner, A. J., Fortgang, R., et al. (2021). A pilot study using frequent inpatient assessments of suicidal thinking to predict short-term postdischarge suicidal behavior. *JAMA Network Open, 4*(3), Article e210591. https://doi.org/10.1001/jamanetworkopen.2021.0591

Wang, S., Yacoby, Y., Pan, W., Bentley, K., Bird, S., Buonopane, R., et al. (2023). Idiographic prediction of suicidal thoughts: Building personalized machine learning models with real-time monitoring data. *PsyArXiv*. https://doi.org/10.31234/osf.io/qh4jk

Webb, C. A., Murray, L., Tierney, A. O., & Gates, K. M. (2023). Dynamic processes in behavioral activation therapy for anhedonic adolescents: Modeling common and patient-specific relations. *Journal of Consulting and Clinical Psychology*. https://doi.org/10.1037/ccp0000830

Weigard, A., Lane, S., Gates, K., & Beltz, A. (2023). The influence of autoregressive relation strength and search strategy on directionality recovery in group iterative multiple model estimation. *Psychological Methods, 28*(2), 379–400. https://doi.org/10.1037/met0000460

Williams, D. R., Mulder, J., Rouder, J. N., & Rast, P. (2021). Beneath the surface: Unearthing within-person variability and mean relations with Bayesian mixed models. *Psychological Methods, 26*(1), 74–89. https://doi.org/10.1037/met0000270

Woods, W. C., Arizmendi, C., Gates, K. M., Stepp, S. D., Pilkonis, P. A., & Wright, A. G. C. (2020). Personalized models of psychopathology as contextualized dynamic processes: An example from individuals with borderline personality disorder. *Journal of Consulting and Clinical Psychology, 88*(3), 240–254. https://doi.org/10.1037/ccp0000472

Wright, A. G. C., Gates, K. M., Arizmendi, C., Lane, S. T., Woods, W. C., & Edershile, E. A. (2019). Focusing personality assessment on the person: Modeling general, shared, and person specific processes in personality and psychopathology. *Psychological Assessment, 31*(4), 502–515. https://doi.org/10.1037/pas0000617

Wright, A. G. C., & Woods, W. C. (2020). Personalized models of psychopathology. *Annual Review of Clinical Psychology, 16*(1), 49–74. https://doi.org/10.1146/annurev-clinpsy-102419-125032

Wright, A. G. C., & Zimmermann, J. (2019). Applied ambulatory assessment: Integrating idiographic and nomothetic principles of measurement. *Psychological Assessment, 31*(12), 1467–1480. https://doi.org/10.1037/pas0000685