PNAS RESEARCH ARTICLE | PSYCHOLOGICAL AND COGNITIVE SCIENCES

Mapping the timescale of suicidal thinking

Daniel D.L. Coppersmith[a,1], Oisín Ryan[b], Rebecca G. Fortgang[a], Alexander J. Millner[a,c], Evan M. Kleiman[d], and Matthew K. Nock[a,c,e]

Edited by Susan Fiske, Princeton University, Princeton, NJ; received September 8, 2022; accepted March 10, 2023

This study aims to identify the timescale of suicidal thinking, leveraging real-time monitoring data and a number of different analytic approaches. Participants were 105 adults with past week suicidal thoughts who completed a 42-d real-time monitoring study (total number of observations = 20,255). Participants completed two forms of real-time assessments: traditional real-time assessments (spaced hours apart each day) and high-frequency assessments (spaced 10 min apart over 1 h). We found that suicidal thinking changes rapidly. Both descriptive statistics and Markov-switching models indicated that elevated states of suicidal thinking lasted on average 1 to 3 h. Individuals exhibited heterogeneity in how often and for how long they reported elevated suicidal thinking, and our analyses suggest that different aspects of suicidal thinking operated on different timescales. Continuous-time autoregressive models suggest that current suicidal intent is predictive of future intent levels for 2 to 3 h, while current suicidal desire is predictive of future suicidal desire levels for 20 h. Multiple models found that elevated suicidal intent has on average shorter duration than elevated suicidal desire. Finally, inferences about the within-person dynamics of suicidal thinking on the basis of statistical modeling were shown to depend on the frequency at which data was sampled. For example, traditional real-time assessments estimated the duration of severe suicidal states of suicidal desire as 9.5 h, whereas the high-frequency assessments shifted the estimated duration to 1.4 h.

suicide | suicidal thinking | ecological momentary assessment

Suicide is a leading cause of death (1). Despite an increase in suicide research over the past few decades (2), there have not been improvements in predicting (2) and preventing (3) suicide. Thus, there is an urgent need for new approaches to understanding suicide (4).

Much of the empirical research on suicide to date has focused on cross-sectional or retrospective studies of the presence or severity of suicidal thoughts/behavior, focusing on the between-person characteristics that distinguish those who do vs. do not experience these outcomes (1, 4, 5). Increasingly, however, researchers have argued that suicidal thoughts/behaviors can best be understood as a process that evolves over time within an individual person (4, 6, 7). By understanding the dynamics of suicidal thinking, that is, how thoughts change over time, we might gain new insights into the mechanisms through which suicidal behavior develops. The recent widespread availability of smartphones has created an opportunity for researchers to begin collecting real-time data on suicidal thoughts (8, 9), providing new insights into the dynamics of suicidal thinking.

One foundational question for understanding the dynamics of suicidal thinking is over what timescale suicidal thoughts evolve (i.e., duration from onset to offset of a given episode of suicidal thinking). The timescale of a process is the cornerstone of any empirical or theoretical approach to understanding it: It determines how researchers should measure the constructs or system of interest and it is the starting point for any mathematical or computational model of underlying mechanisms of action. If the timescale of a process is slow, it is not necessary to measure it as frequently as a faster process (10): It would be sufficient to measure the physical growth of a child once a day, but to capture the reaction of heart rate to caffeine intake, a one-day gap in measurements would be insufficient. Unfortunately, very little is known about the timescale of suicidal processes because we have not studied how suicidal thoughts or behaviors change over minutes, hours, or days (4, 11). As a result, researchers and clinicians have little guidance on how often to measure suicidal thinking or how to appropriately formalize theories on suicidal behavior.

In this paper, we aim to gain insights into the timescale of suicidal thoughts using an ecological momentary assessment dataset consisting of a mix of a) observations taken at intervals of several hours apart and b) burst observations spaced at a higher frequency of 10 min apart. The design of the study is depicted in Fig. 1. We quantify the timescale of suicidal thoughts using three approaches. First, we use descriptive statistics to examine any systematic differences in suicidal desire and intent ratings resulting from high- and low-frequency measurements. Second, we use summary statistics to quantify the rate at

Significance

Suicide is one of the most devastating aspects of human nature and has puzzled scholars for thousands of years. Most suicide research to date has focused on establishing the prevalence and predictors of the presence or severity of suicidal thoughts/behaviors. Surprisingly, little research has documented the fundamental properties of suicidal thoughts/behaviors, such as when someone has a suicidal thought, how long do such thoughts last? Documenting the basic properties of a phenomenon is necessary to understand, study, and treat it. This study found that elevated states of suicidal thinking lasted on average 1 to 3 h. These results provide the most detailed characterization to date of the temporal dynamics of suicidal thinking.

Author affiliations: [a]Department of Psychology, Harvard University, Cambridge, MA 02138; [b]Department of Methodology and Statistics, Utrecht University, 3508 TC Utrecht, The Netherlands; [c]Mental Health Research, Franciscan Children's, Brighton, MA 02135; [d]Department of Psychology, Rutgers, The State University of New Jersey, Piscataway, NJ 08854; and [e]Department of Psychiatry, Massachusetts General Hospital, Boston, MA 02114

Author contributions: D.D.L.C., R.G.F., A.J.M., E.M.K., and M.K.N. designed research; D.D.L.C., R.G.F., E.M.K., and M.K.N. performed research; O.R. contributed new analytic tools; D.D.L.C. and O.R. analyzed data; and D.D.L.C., O.R., R.G.F., A.J.M., E.M.K., and M.K.N. wrote the paper.

Competing interest statement: M.K.N. receives publication royalties from Macmillan, Pearson, and UpToDate. He has been a paid consultant in the past year for Cerebral, Compass Pathfinder, and for a legal case regarding a death by suicide. He is an unpaid scientific advisor for Empatica, Koko, and TalkLife. These roles are not perceived as creating conflicts of interests but are reported for transparency.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: dcoppersmith@g.harvard.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2215434120/-/DCSupplemental.

Published April 18, 2023.

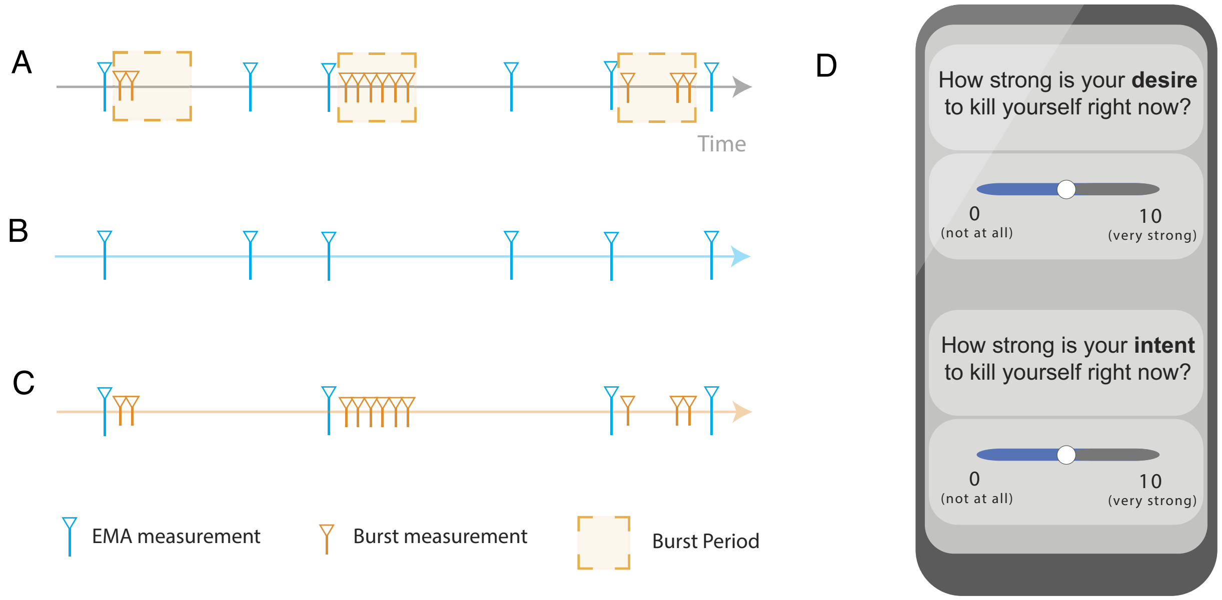

Fig. 1. Overview of sampling design. (*A*) combined EMA and burst measurements; (*B*) EMA-only dataset; (*C*) Burst-only (short) dataset; and (*D*) Questions asked about suicidal thinking during real-time measurements

which we observe changes in reported suicidal desire and intent. Finally, we use continuous-time models (12, 13) to model the dynamics of suicidal thinking.

Results

Descriptive Statistics. The 105 participants of the study completed a total of 20,255 surveys. The number of observations per person ranged from 5 to 456 (mean = 193, IQR = 58 to 296). The compliance percentage per person ranged from 1% to 92% (mean = 39%, IQR = 12 to 60%). The burst compliance is linked to the EMA compliance in that the burst surveys would not fire if the participant did not first complete an EMA survey. We provide a conservative estimate of the burst survey compliance by estimating it based on the maximum number of burst surveys a participant could complete in the study. The mean compliance for the EMA surveys (44%) was higher than the mean compliance for the burst surveys (35%). The mean length of the observation period, that is, the time elapsed between the first and last observations, was 33.64 d (min = 2, max =41.62, IQR = 30.03 to 41.08).

We assess whether measurements taken at different frequencies (burst vs. EMA) are capturing the same underlying processes by examining the (within-person, over time) means and variances of desire and intent. Fig. 2 panels *A* and *B* show that the distributions of desire and intent ratings are highly similar across burst and EMA measurements. In panel *C*, we see that individual's EMA-mean and burst-mean are highly correlated (r = 0.94 for desire and r = 0.97 for intent) and in panel c that the within-person SDs for both measurement types are highly correlated (r = 0.74 desire and r = 0.88 for intent). The similarity of these characteristics across measurement types indicates that participants do not appear to be responding to EMA and burst measurements in a systematically different manner. We recorded 1,213 high-risk observations (6% of all observations), spread across 31 participants (see *Methods and Materials* for the definition of high-risk observations). Then, 544 of these high-risk observations were captured by traditional EMA samples (544/9,755 = 5.6% of all EMA measurements were high risk). The remaining 669 were captured by burst surveys (669/10,500 = 6.4% of all burst measurements were high risk). This suggests that high-risk moments were slightly more likely to be observed with burst measurements than with traditional EMA. To account for the fact that high-risk states may last relatively long, we computed the number of unique high-risk moments captured by burst measurements (high-risk moments reported during burst in which the preceding EMA measurement was not high risk). In total, 74 unique high-risk observations were captured by burst measurements. In terms of individuals, in total six individuals recorded high-risk moments in burst measurements while never recording a high-risk moment in an EMA measurement. In other words, with burst measurements, we identified 19% more participants with a high-risk moment.

For the majority of participants, the most common response to the self-report items, i.e., the modal response, was 0 (not at all): (59/105) 56.2% of participants had a modal response of 0 to the desire item and (78/105) 74.3% had a modal response of 0 to the intent item. We can consider the most common response to be an indicator of the resting state of desire and intent for a given individual (14). All participants who had a mode of 0 on desire also had a mode of 0 on intent.

The degree of variability in responses also differed across individuals and items. Since the items are measured using a scale limited to integers between 0 and 10, we can quantify variability in terms of how often a participant reacts to a measurement prompt by using a response category which is equal to their modal or resting state value. We quantify this using a statistic we will refer to as p_{mode}, with $p_{mode} = 1$ indicating that an individual responded with the same answer category to every prompt and therefore shows no variability. We see that typically around 50 percent of an individual's responses to the desire item were different from that individual's mode, but we also observe a large degree of variability in this response pattern across individuals (median $p_{mode} = 0.50$, IQR = 0.30 to 0.83), with one individual reporting no variability at all. Variability on the intent item was even lower (median $p_{mode} = 0.91$, IQR = 0.49 to 0.99), with 15 individuals showing no variability at all in reported suicidal intent. All individuals who exhibited zero variability in responses had a resting state of zero, and there was a high positive correlation between the variability metrics of desire and intent across participants ($r = 0.621$).

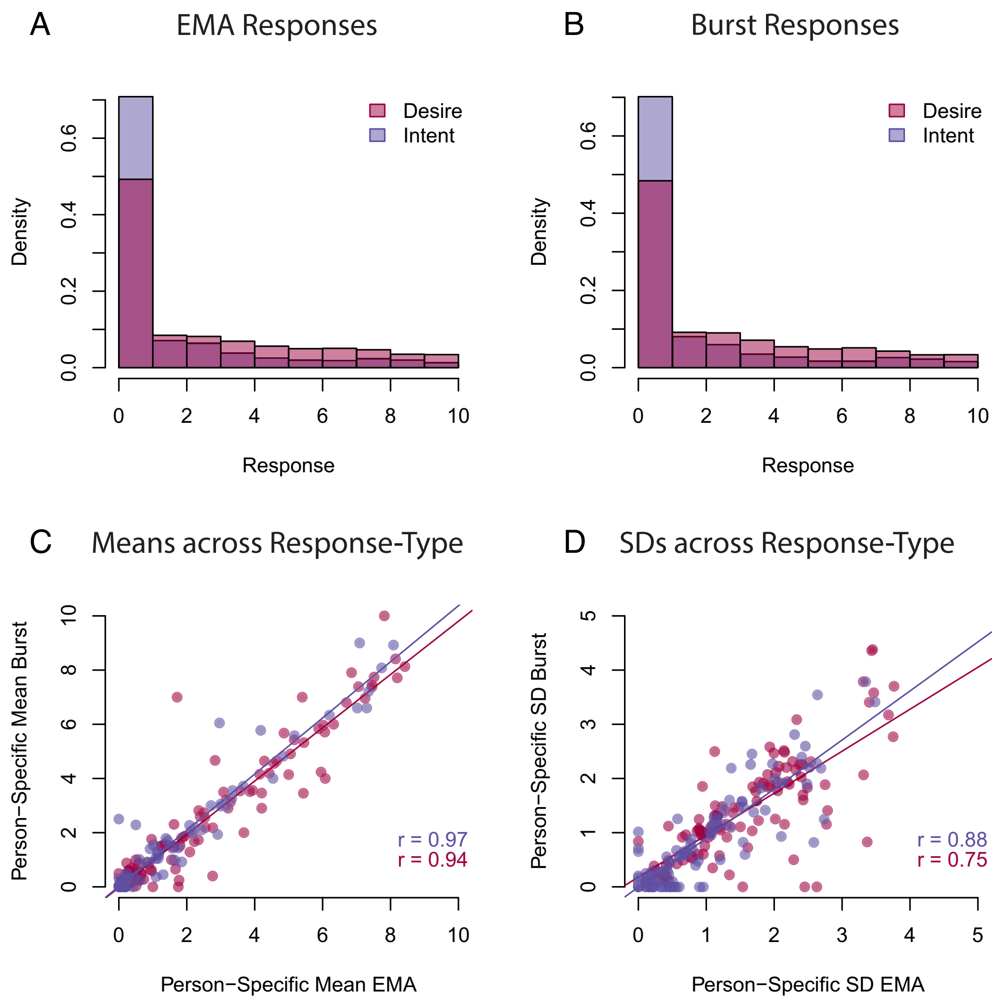

Fig. 2. Descriptive statistics across different types of real-time measures of suicidal thinking. (*A*) The distribution of EMA responses for desire and intent; (*B*) The distribution of burst responses for desire and intent; (*C*) The means of desire and intent across response-type; (*D*) The standard deviations of desire and intent across response-type.

Timescale Summary Statistics. Visual inspection of time-series with low and high variability, respectively, suggest a further qualitative difference between the observed time-series. As we can see from the example shown in Fig. 3*A*, empirical time-series which exhibited low-variability also appeared to exhibit an episodic pattern of dynamics: In these time-series, responses consist of long sequences of zeros, interspersed by sequences of nonzero responses, which we can interpret as indicating an episode of heightened suicidal thinking lasting for a relatively short period of time. In contrast, time-series which have relatively high variability, as the example shown in Fig. 3*B*, do not appear to exhibit episodes of heightened responses but instead vary in a more continuous fashion from measurement occasion to measurement occasion.

Using the p_{mode} statistics described above, we categorize observed time-series as episodic (high p_{mode}) or nonepisodic (low p_{mode}) in nature (see *Methods and Materials* for details). In total, we counted 152 episodes of elevated desire across 27 participants with a median estimated duration of 1.96 h (IQR = 0.35 to 3.02). We count 81 episodes of elevated Intent events spread across 43 individuals, with a median duration of 1.0915 h (IQR = 0.21 to 2.70). This shows that for participants who report elevated desire and intent as rare episodes or events during our observation window, the event itself is typically quite short lived—lasting from a number of minutes up to typically around 1 to 2 h, with only rare exceptions of much longer-lasting episodes.

For nonepisodic time-series, we calculate how often pairs of consecutive measurements show variability and examine how this variability depends on the amount of time between those measurement occasions. The results for desire and intent are shown in Fig. 4. For desire, we used data from n = 85 individuals and for intent n = 58 individuals (see *SI Appendix* for more details on data filtering). The proportion of change increases as the duration of timescale increases, which is to be expected. There was substantial variation (on average 42.2% for desire, 44.2% for intent) in the shortest timescale (less than 45 min). It is important to note that Fig. 4 highlights individual differences in both the amount of variation and how variation changes across timescales. In the *SI Appendix*, we show that using a more conservative p_{mode} cut-off score to categorize observations as episodic or nonepisodic leads to similar results.

Dynamics of Suicidal Thinking. The Fig. 5 panel (*A*) shows the continuous-time vector autoregressive (CT-VAR) model fixed effects (drift matrix) estimates along with their 95% credible intervals (13). The parameters should be interpreted as estimates of the average moment-to-moment dynamics across individuals. We see that the model estimates a positive (i.e., activating) effect from desire to intent but no reciprocal effect from intent to desire. The interpretation of these parameters is that elevations in levels of desire are predicted to cooccur with future elevations in intent but that elevations in Intent do not predict changes in the level of

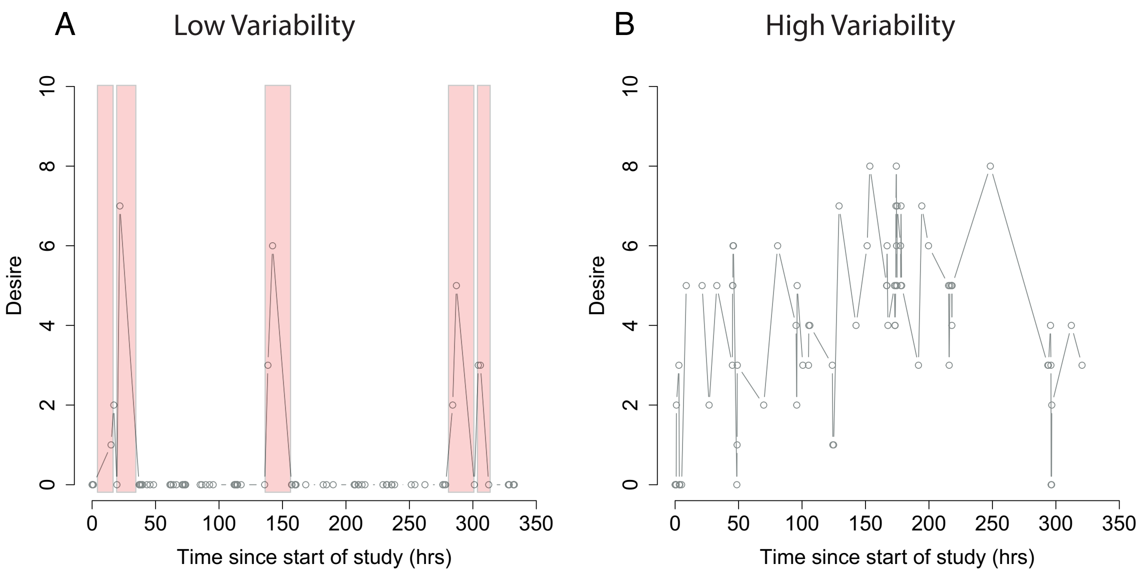

Fig. 3. In panels (*A*) and (*B*), we show a time-series depicting the first two weeks' worth of responses on the Desire variable for two different participants. In panel (*A*), the time-series is taken from a participant with low variability, indicated by $p_{mode} = 0.91$. The red boxes indicate periods of consecutive measurement occasions in which the participant gave a nonzero response on the Desire item, referred to in the main text as episodes of elevated desire. In panel (*B*), the time-series is taken from a participant with high variability, indicated by $p_{mode} = 0.22$

desire. The self-loops depicted in the graph represent the estimated effect of each variable on its own rate of change. These parameters are negative and so can be interpreted as self-regulating effects or dampening effects of desire and intent on themselves over time: The further away from equilibrium the process finds itself, the quicker the process moves toward equilibrium (15).* The more strongly negative self-effect of intent should be interpreted as reflecting that elevated levels of intent are more quickly regulated away than elevated levels of desire.

Panels (*B*) and (*C*) of Fig. 5 show how the moment-to-moment dynamics can be used to understand the implied timescale of the process. Panel (*B*) shows how, according to the parameter estimates, current observations of desire and intent are expected to be predictive of each other's future values, as a function of the time interval between observations. We see that desire is predictive of itself and intent for a period of around 20 h; the relationship between current desire and future intent is strongest at a time interval of around 2 to 3 h. We can also see that Intent is weakly predictive of desire over any time interval, and current levels of intent no longer have any predictive value for future Intent levels after 2 to 3 h. Another way to understand the timescale of change implied by the model parameters is the impulse response function (IRF), that is, the predicted trajectory of the system over time given an impulse. Panel (*C*) shows the predicted IRF given an impulse to increase levels of Desire. Again, we see the model predicts that Intent will increase over a period of 2 to 3 h before both desire and intent return to baseline over a period of 20 h.

The Fig. 6 panels (*A*) and (*B*) show the CT Markov switching model estimates for the desire and intent variables, respectively.

*The self-loops depict the diagonal elements of the drift matrix, sometimes referred to as autoeffects. Discrete-time autoregressive effects between 0 and 1 often interpreted as inertia parameters can be interpreted in the same way as continuous-time autoeffects running from 0 to negative infinity.

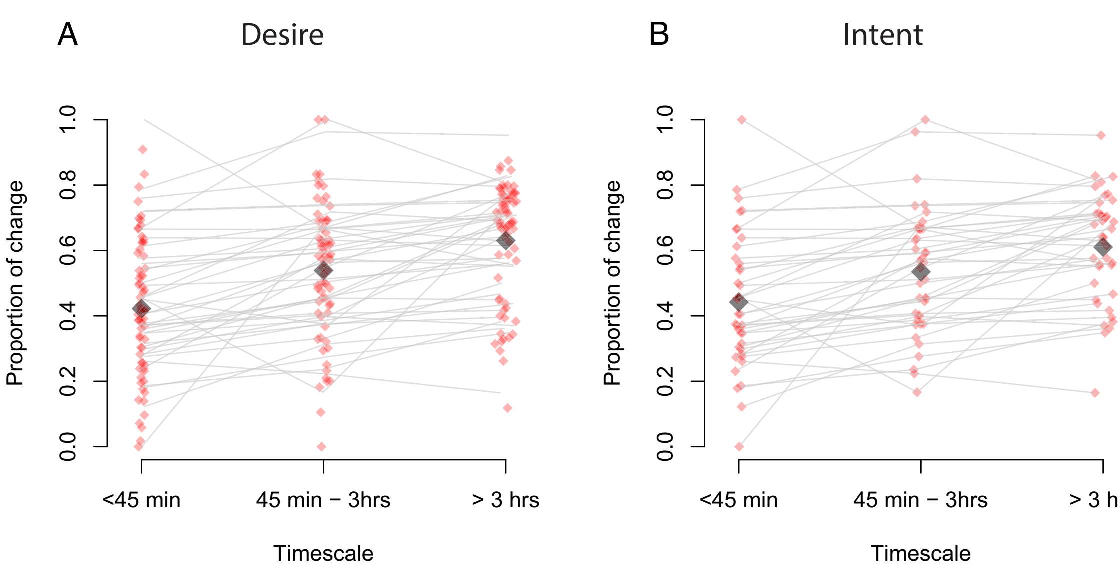

Fig. 4. Proportion of consecutive observations that show variation as a function of time between observations for desire (*A*) and intent (*B*). Each red dot represents an individual participant, with the light gray lines connecting values of the same participant across timescales. The gray diamond represents the mean proportion in a given timescale (Desire = {0.422, 0.539, 0.631}, Intent = {0.442, 0.535, 0.610})

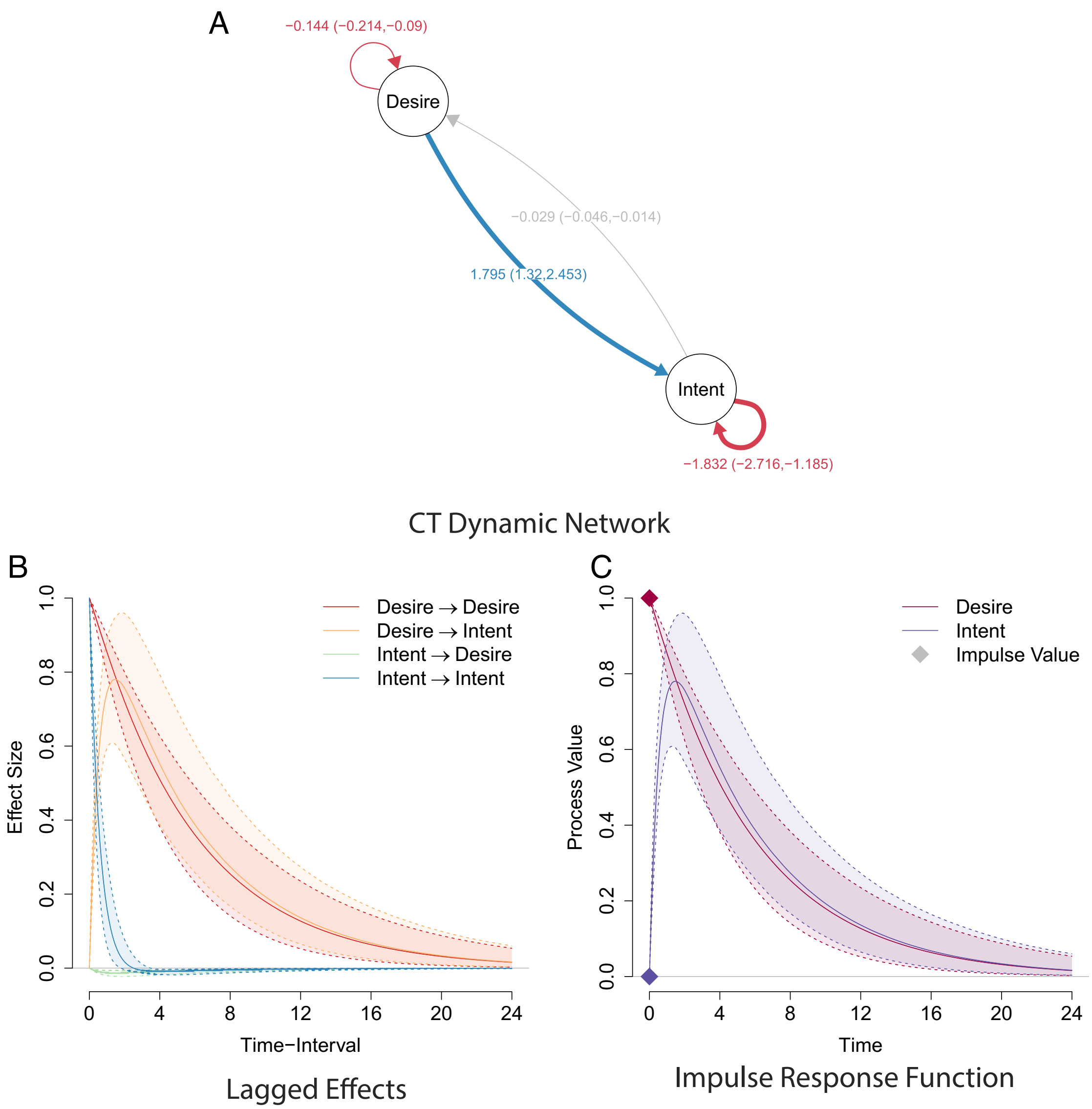

Fig. 5. Continuous-time vector autoregression results. Panel (*A*) depicts the estimated drift matrix fixed effects, with CIs, as a network. Panel (*B*) shows how the model-implied lagged regression coefficients are dependent on the time interval between measurements . Shaded lines represent 95% credible intervals. Panel (*C*) shows the model-implied Impulse Response Function, that is, how the model predicts the values of Desire and Intent to change over time given an impulse value of (Desire = 1 and Intent = 0) as indicated by the filled diamond.

For the sake of interpretability, we use the estimated intensity matrix (*SI Appendix, S.2*) to derive, on the left-hand-side of each figure, the estimated transition probability matrix over a time interval of 1 h, that is, the probability of either staying in the same state (self-loops) or transitioning to a different state at an assessment taken 1 h later. For both desire and intent, we observe that, at 1-h intervals, participants are most likely to stay in their current state and that the probability of staying in the zero and low states are largest, reflecting that overall participants seem to be relatively stable in reporting no or low levels of suicidal thinking. On the right-hand-side of each panel, we show what are known as the estimated sojourn times for each state, that is, how long on average a participant is predicted to stay in one state. The sojourn times thus represent a direct model-based estimate of the timescale of suicidal desire and intent in terms of estimated state durations. We see that on average, zero levels of desire are estimated to last 12.90 h [CI: 11.26, 14.79]; low levels of desire are estimated to last 5.84 h [CI: 5.33, 6.42]; moderate levels 1.76 h [CI: 1.59, 1.96]; and severe levels 2.83 h [CI: 2.43, 3.29]. We see a similar pattern for intent, with zero levels estimated to last 8.41 h [CI: 7.46, 9.49], low levels estimated to last 3.87 h [CI: 3.52, 4.26], moderate levels 1.05 h [CI: 0.93,1.19], and high levels 2.92 h [CI: 2.36, 3.62].

As a final analysis, we estimated both the CT-VAR and CT-Markov model on subsets of our data consisting of only regular EMA measurements and short time interval burst measurements, respectively (see Fig. 1 panels *B* and *C*). Fig. 7 shows the CT-VAR parameter estimates across both datasets. Comparing panels *A* and *B*, we see many qualitative similarities in parameter estimates. The self-effect of intent in both cases is higher than that of desire; desire has a positive cross-effect on intent, although the weakly negative effect of intent on desire is considered "significant" in the EMA-dataset, in that the upper end of the credible interval does not cross zero. The parameter estimates themselves, however, are quite different and yield somewhat different interpretations of the timescale of suicidal thinking. Based only on the EMA measurements, we would derive that elevated levels of intent return to baseline much more slowly than we would conclude based on the full dataset or the short-time-interval subset. This can be seen in panels (*C*) where the self-effect of intent reaches zero after 8 h, and in panel (*E*) where, according to the IRF, the effect of an impulse on intent is still present

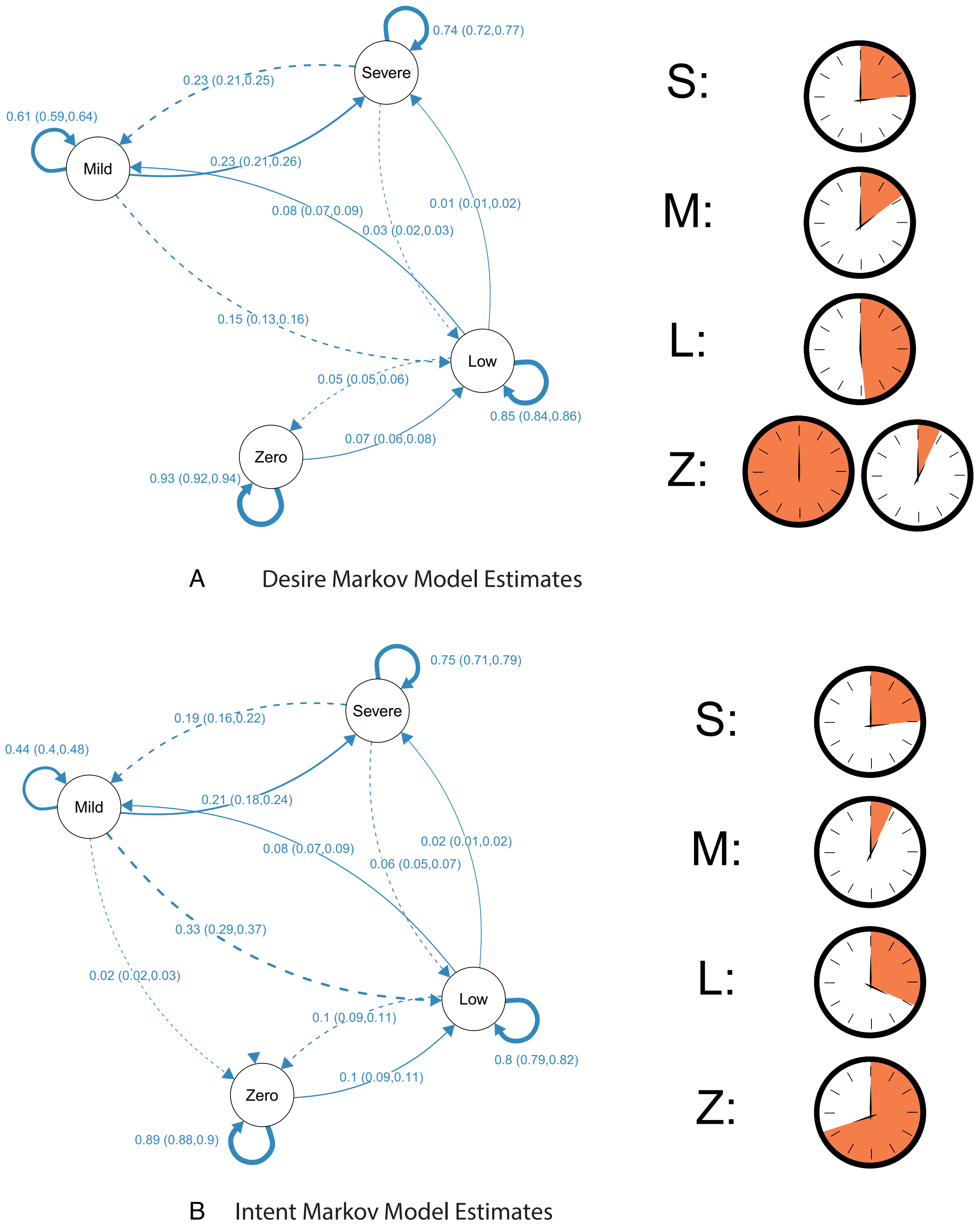

Fig. 6. Transition Probabilities (at a 1-h interval, *Left*) and Sojurn Times (*Right*) from the Continuous Time Markov Models. Clocks represent sojurn time duration in blocks and fractions of 12 h. In the networks, transition probabilities point estimates that are smaller than or equal to 0.001 in absolute value are omitted. (*A*) Desire Markov model estimates; (*B*) Intent Markov model estimates.

after 24 h. Fig. 8 shows the CT-Markov parameter estimates across both datasets. We again see a qualitatively similar pattern of results but that both datasets yield different quantitative estimates of the timescale of change of suicidal thinking. Across both desire and intent, the probability of transitioning was higher in the high-frequency dataset. For example, the probability of transitioning from a mild state to a severe state for desire in the EMA dataset is 9.2%, whereas in the high-frequency dataset, the probability for desire is 29%. For both desire and intent, the average sojourn times for mild and severe states were shorter in the high-frequency dataset

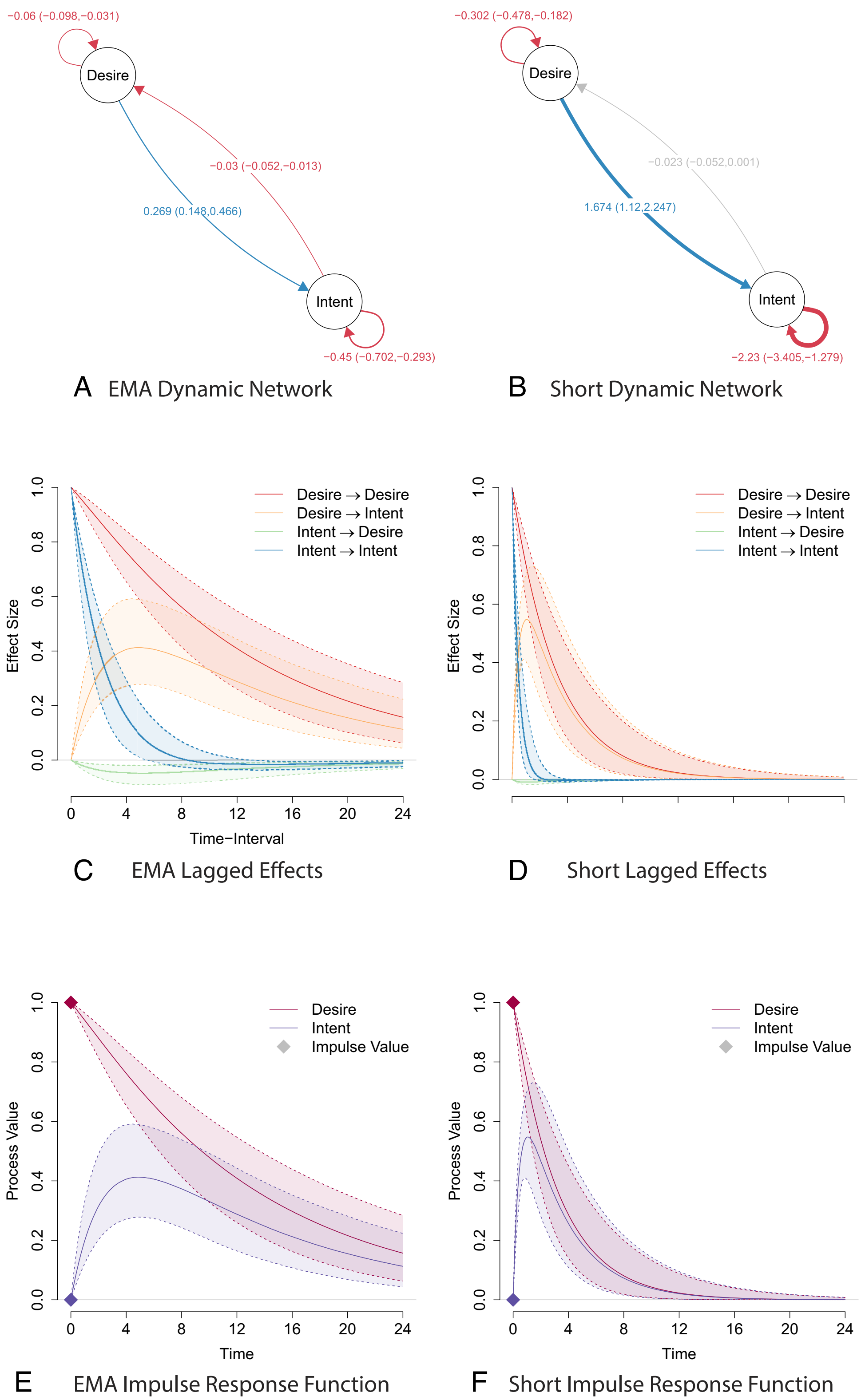

Fig. 7. Continuous-time vector autoregression results across EMA and short datasets. (*A*) EMA dynamic network; (*B*) Short dynamic network; (*C*) EMA lagged effects; (*D*) Short lagged effects; (*E*) EMA impulse response function; (*F*) Short impulse response function.

than the EMA dataset. In the EMA datasets, the average sojourn times for mild and severe states was 5.4 and 9.5 h (desire) and 3.8 and 9.3 h (intent). In the high-frequency datasets, the average sojourn times for mild and severe states were merely 0.84 and 1.42 h (desire) and 0.47 and 1.25 h (intent). The average sojourn times for each state are provided in the *SI Appendix*.

Discussion

This study represents the highest-resolution examination of the temporal dynamics of suicidal thinking to date. There were three key findings. First, suicidal thinking changes rapidly. Second, different aspects of suicidal thinking (i.e., suicidal desire and intent)

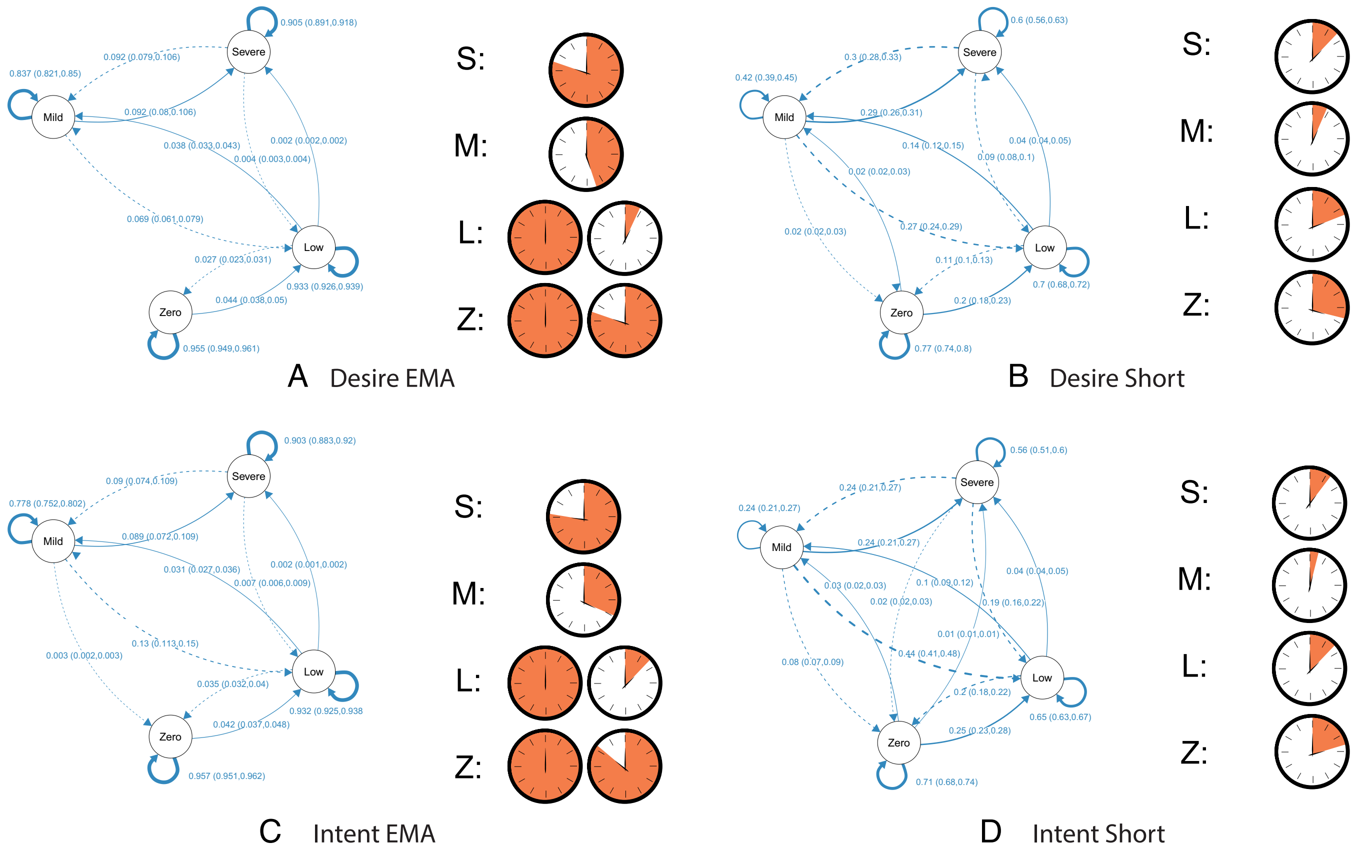

Fig. 8. Transition Probabilities (at a 1-h interval) and sojurn times from continuous time Markov models across EMA and short datasets. (*A*) Desire EMA Markov model estimates; (*B*) Desire short Markov model estimates; (*C*) Intent EMA Markov model estimates; (*D*) Intent short Markov model estimates. Note: clocks represent sojurn time duration

operate on different timescales. Third, the within-person dynamics of suicidal thinking captured by state-of-the-art statistical models differ depending on whether we use higher or lower frequency sampling. Each of these findings warrants additional comment.

Across timescale summary statistics and multiple statistical models, we found evidence that suicidal thinking changes rapidly. Suicidal thinking returned to baseline within several hours and elevated suicidal desire led to elevated suicidal intent within minutes to hours. These findings provide empirical support to theoretical work that conceptualizes of suicide as a dynamic system (6) and preliminary descriptive research on the dynamics of suicidal thinking (16). Converging lines of research (17–19) support conceptualizing of suicide risk as a dynamical system that unfolds over short timescales. This dynamic conceptualization of suicide should inform future theories (4), assessment (18), and interventions (20, 21).

From this study, we learned several things about different aspects of suicidal thinking. First, elevated levels of suicidal desire appear to last longer than elevated levels of suicidal intent. Second, suicidal desire appears to precede or predict later suicidal intent. Prior research has often either measured suicidal thinking with a single-item assessment (22, 23) or collapsed multiple aspects into a single sum score of suicidal thinking (9, 24). Our findings suggest taking such an approach would be problematic because different dimensions of suicidal thinking have unique dynamics. Through providing greater specificity in characterizing different dimensions of suicidal thinking, our findings support conceptualizing of suicidal thinking as a multidimensional phenomenon. This more precise understanding of the different dimensions has implications for both assessment and intervention. For example, we characterized the duration from an instance of elevated suicidal desire to suicidal intent. This time between desire and intent represents a potential window for intervention (20) to reduce risk.

Finally, we showed how sampling frequency matters for characterizing suicidal thinking. If one were only interested in extracting features (e.g., the mean) from the time-series of suicidal thinking (18), lower-frequency sampling would show no difference from higher-frequency sampling. The proposed promise, however, of real-time monitoring research has been in capturing dynamics as they unfold (8, 25). Our results suggest that accurately characterizing within-person dynamics of suicidal thinking requires higher-frequency sampling (i.e., minutes apart). State-of-the-art statistical models likely cannot make up for under-sampled data when the underlying dynamic model is not known or correctly specified (26). While continuous-time models have numerous advantages when compared to discrete-time models in terms of correcting for different sampling frequencies (12), in our analysis, high-frequency measurements still yielded different quantitative estimates of the timescale of suicidal thinking than lower-frequency EMA measurements. There may be different benefits and costs to different sampling approaches. For example, a risk of using classic real-time monitoring designs (i.e., hours apart) is that one may underestimate the speed of the system because one cannot infer shorter lags (e.g., 30 min) from longer lags (e.g., 6 h). A cost of higher-frequency sampling is that it may be burdensome on participants and results in lower overall compliance (27). The implication of these findings is not that all psychological constructs need to be measured in brief (e.g., 10 min) intervals but rather highlights the crucial importance of sampling frequency as a design decision. For example, one may wish to use high-frequency sampling over a brief (e.g., 1 wk) high-risk period and more classic designs to study risk over longer timescales.

While the current study has several strengths such as the sampling design, there are multiple limitations that require discussion. First, a concern of high-frequency assessments is reactivity to the assessments, such as that monitoring changes the process under observation. While analyses from this overall project (28) and other projects (29) suggest that frequently assessing suicidal thinking does not increase suicidal thinking, it could have impacted the data in some other way. Second, it is possible that part of the variation in responses across measurements occasions is attributable to the presence of measurement error, which may also partially explain the differences in continuous-time model estimates across different subsets of the data. Unfortunately, there may be no straightforward remedy for this problem in high sampling frequency designs. Asking multiple versions of the same question may not be practically feasible in such a design due to participant burden, while averaging over different suicidal thinking items may lead to erroneous conclusions if they represent distinct processes with different timescales. Third, for continuous-time models, we were only able to use the subset of the participants with high compliance and some variability in their suicidal thinking. This limits the generality of the findings. Fourth, for the Markov models, we arbitrarily categorized suicidal thinking into three states, an approach which could be viewed as correcting for variation due to measurement error if these reflect true states. In future work, hidden Markov models (30) could be applied for a data-driven approach to identifying states of interest. Fifth, while we used two of the most widely studied continuous-time models to study the dynamics of suicidal thinking, these models are relatively simple and so model misspecification is probable: The dynamics governing intent and desire may not be well approximated by linear and first-order dynamic models. Although the Markov models and the mode-based descriptive statistics took different approaches to characterizing the timescale of suicidal thinking, both approaches can be interpreted as indicating that elevated levels of suicidal thinking last on average 1 to 3 h. More exploratory tools (31) are needed for data sampled at uneven time intervals (26). Sixth, the current paper only used a sample of adults recruited online. It is currently unknown if these findings on the dynamics of suicidal thinking generalize to other populations, such as adolescents with suicidal thoughts or adults hospitalized for suicide risk. Seventh, the compliance for the real-time surveys was relatively low compared to traditional real-time monitoring studies. It is unknown, however, what are realistic expectations of compliance with high-frequency sampling.

Finally, the current study only measured two aspects of suicidal thinking and did not measure suicidal behavior. It is currently unknown how these very short dynamics relate to suicidal behavior.

There are several future directions that build upon the current study. First, one could replicate this study with a clinical sample, such as psychiatric inpatients, to test the generalizability of the findings. Second, one could use higher-density sampling over a longer interval than 1 h. For example, one could try to sample every 30 min over the course of a day with a small number of questions of suicidal thinking. This would allow for building a continuous time model for a 24-h period. Third, one could track suicidal behavior after the real-time monitoring period to understand if higher sampling of suicidal thinking improves the prediction of future suicidal behavior (18). This would be insightful for further understanding the value gained by more frequent assessments. Finally, the ultimate aim of the current study was to richly characterize the dynamics of the phenomenon of suicidal thinking. Future theoretical work could attempt to integrate these dynamics into a formal theory of suicide (4, 32). A valuable test of a formal theory of suicide would be evaluating whether it could recover the dynamics identified in the current study.

Suicide has historically represented one of the most difficult topics to study. The current study highlights how we can use new forms of data collection and statistical models to zoom the microscope in on this perplexing phenomenon. We have provided insights into suicidal thinking and hope this work catalyzes conceptualizing of suicide risk as a dynamical process that unfolds over short windows of time.

Methods and Materials

Participants. Participants were 105 adults recruited through online advertisements. The average age of participants in this sample was 29.22 y (SD = 9.10, range = 18 to 60). Most participants (68.57%) reported being assigned female sex at birth (n = 72), 30.48% reported being assigned male sex at birth (n = 32), and 0.95% did not report assigned sex at birth (n = 1). For gender, 56.19% identified as cisgender female (n = 59), 33.33% identified as cisgender male (n = 35), 1.90% identified as transgender (n = 2), 7.62% identified as nonbinary (n = 8), and 0.95% did not report (n = 1). Regarding racial identity, 21.90% of participants identified as multiracial (n = 23), 8.57% as Black (n = 9), 4.76% as Asian (n = 5), 57.14% as White (n = 60), and 1% as Middle Eastern (n = 1). Race was unknown for 6.66% of participants (n = 7). For ethnicity, 15.24% of participants identified Latinx (n = 16). The median number of lifetime days with suicidal thoughts was 1,825 (range = 30 to 8,000 d). More than two-thirds of participants (65.74%) reported a prior suicide attempt (n = 71).

Participant Recruitment. Participants were recruited online (via Reddit and Craigslist). Participants completed an eligibility screener that assessed self-injurious thoughts and behaviors. Inclusion criteria were active suicidal thoughts in the past week, fluency in English, >17 y of age, and regular access to a smartphone. An exclusion criterion was living in Europe due to GDPR restrictions, failing a CAPTCHA test, or providing inconsistent or illogical responses. Participants did not have to provide their name, address, or phone number. Participants only provided an email address for communication about study procedures and sending payment. One benefit of relatively anonymous data collection is that anonymous sampling is associated with higher rates of endorsement for suicidal thinking (33, 34). The higher rates of endorsement may be due to participants' concerns about reactions to accurately disclosing suicidal thoughts (35, 36).

A total of 8,035 individuals completed the recruitment screener; 279 qualified and were emailed a baseline assessment; 161 completed it but 30 were removed for denying active suicidal thoughts or responding inconsistently; 131 qualified for the study; and 115 downloaded the application. The first five participants were pilot subjects for feasibility testing and were excluded from analyses and another five failed to complete a single momentary assessment (n = 3) or responded inconsistently (n = 2). A total of 105 participants were included in the final sample.

Procedure. The study procedure consisted of a baseline survey, a 42-d real-time monitoring period, and 4-wk follow-up survey. Participants were compensated with Amazon gift cards. Participants were paid $5 for the baseline survey and $5 for the follow-up survey. Participants were paid $0.25 for each completed survey and a $1 bonus per for any day where they completed at least five surveys. The maximum amount participants could earn in the study was $190. All study procedures were approved by the Harvard University-Area Institutional Review Board (IRB# 19-1819; "High-Resolution Real-Time Capture of Suicidal Thoughts and Urges"). Informed consent was obtained from all participants.

Real-Time Monitoring. We used Metricwire to collect real-timed monitoring data. For the real-time monitoring period, participants downloaded the Metricwire app that sent them three types of surveys over a 6-wk period daily surveys (one time per day), EMA surveys (five times per day), and burst surveys (6 times/h, 2 per d, 4 d/wk). In this paper, we focus on the EMA surveys and burst surveys (Fig. 1). EMA surveys were sent five times per day at least 90 min apart between participant custom wake/sleep times. EMA surveys stayed open for 1 h. Burst surveys were sent six times per hour (episode) with the maximum of two episodes per day and 4 d per wk. The burst surveys were sent immediately following the completion of a longer EMA (Fig. 1). Each burst survey stayed open for 5 min. Following recommendations for anonymous real-time studies (37), automated messages were used for suicide risk monitoring. Specifically, in each EMA survey, participants were provided with a pop-up message in the Metricwire app providing resources for treatment and safety (e.g., suicide prevention hotlines).

Real-Time Monitoring Measures. Suicidal thoughts were measured with two items: desire to kill self and intent to kill self "right now" (Fig. 1*B*). Both items were rated on a 0- (not at all) to 10- (very strong) Likert scale. Desire was defined for participants as how much do you want to kill yourself. Intent was defined as to what extent are you actually going to kill yourself. These items were selected based on the Beck Scale for Suicidal Ideation (38), a widely used measure of suicidal thinking. Questions pertaining to desire and intent are similar to momentary measures of suicidal thinking used in other real-time monitoring studies (9, 16) and have shown predictive validity of suicidal behavior (18).

Analytic Approach. All analyses were conducted in R version 4.10. All data analysis codes are available at https://github.com/ryanoisin/TimescaleSuicidalThinking. The data and materials that support the findings of this study are available from the corresponding author on request. We used three different strategies to leverage our unique sampling design to gain insight into the timescale of change of suicidal thinking.

First, we examined descriptive and time-invariant properties of our variables, such as the within-person mean, variance and mode, and the number of high-risk responses captured, with the latter defined as ratings of 8, 9, or 10 on the suicidal intent item based on previous research (37). We assessed whether measurements taken at different frequencies capture fundamentally different processes on the aggregate level by examining whether these characteristics differed between high-frequency (from the burst design, every 10 min) and low-frequency (from the EMA design, approximately 3 to 12 h apart) measures.

Second, we used summary statistics to quantify the rate at which we observe changes in self-reported desire to kill self and intent to kill self within individuals over time. The large degree of heterogeneity that was observed between individuals in terms of response patterns over time motivated the use of two different analytic approaches to achieve this aim. Through visual inspection of individual time-series (Fig. 3), we identified that low-variability time-series appeared to vary in an episodic manner over time, consisting of distinct periods of elevated responses, while higher-variability time-series did not exhibit episodic patterns, instead varying in a more continuous fashion over time.

Based on this observation, we categorize empirical time-series as episodic if $P_{mode} < 0.8$. For episodic time-series, we quantify the rate of change in the time-series by calculating the frequency and duration of episodes of elevated suicidal thinking. We achieve this by categorizing sequences of consecutive observations that show elevated desire or intent per person as an episode; we then calculate the duration of elevated desire/intent by observing how much time elapses before the next occasion where desire or intent are back to their resting state. As such, we can consider the duration estimated in this way as an upper bound of the true episode length (as participants may have returned to "normal" before the next observation). In our analysis, we omit duration estimates which include night gaps (observations which are elevated before bed, but the next day back to normal).

For nonepisodic time-series ($P_{mode} \geq 0.8$), we quantify the rate of change by calculating how often pairs of consecutive measurements differ in value from one-another and by examining how this variation depends on the length of the time interval between those measurement occasions. For each individual, we bin pairs of consecutive observations according to how far apart in time they are spaced. We then calculate the degree of variation at different timescales by calculating how often these pairs of consecutive observations differ in value. To facilitate this analysis, we create three timescale bins: less than 45 min, 45 min to 3 h, and greater than 3 h. These bins were chosen such that, as much as possible, all individuals have at least two pairs of observations within each bin. We calculate the rate-of-change for each individual in each bin and also calculate averages of these within-person statistics across individuals. This allows us to gain direct insight into the timescale of change of these variables since we would expect that processes which change quickly will show a larger degree of variation over short time intervals than processes that change slowly over time. Finally, we conducted a sensitivity analysis, presented in the *SI Appendix*, to examine the effect of choosing a more conservative cut-off of $P_{mode} < 0.95$ vs. $P_{mode} \geq 0.95$ on both sets of summary statistics.

Finally, we used continuous-time (CT) models (39) to estimate the moment-to-moment dynamics of suicidal desire and intent. CT approaches allow users to estimate models from data collected at different intervals by explicitly modeling lagged relationships (e.g., the predictive relationship between the current value of Desire and the value of Desire at the next measurement occasion) as dependent on the time interval (that is, the amount of time that elapses) between measurement occasions. In this way, CT models can be seen as a more appropriate alterative to discrete-time models such as vector autoregressive or Markov-switching models more frequently used in psychological settings (40, 41), since the latter assume evenly spaced data and/or that lagged relationships do not depend on the spacing of measurements, assumptions which are almost always violated in real-time monitoring settings (39). This feature also means the estimates of CT models can be used in a straightforward way to compare model parameters from datasets which contain different sampling schemes (12).

In the current paper, we fit CT versions of both the vector auto regressive model (CT-VAR; (12, 13) and the Markov multistate models (CT-Markov; 42) to the empirical data. The CT-VAR models desire and intent as continuous-valued processes, which influence each other over time through a linear system of first-order differential equations. From a qualitative perspective, they model processes which fluctuate around a stable equilibrium: External shocks (random noise) push the processes away from equilibrium, and the underlying lagged dependencies pull the process back to equilibrium over time. The CT-Markov model, in contrast, treats these processes as consisting of discrete states, modeling the probability of changing from one state to another at the next measurement occasion given the current state. To enable us to use this model, we recoded the 11-point scales into four states: a zero state (0), low (1 to 4), mild (5 to 7), and severe (8 to 10).

Both model types use information about the time interval between measurement occasions in order to estimate moment-to-moment dynamic relationships (known as the drift matrix for the CT-VAR and intensity matrix for the CT-Markov). These in turn can be used to model how lagged regression coefficients and state-switching probabilities, respectively, depend in a nonlinear way on the time interval between measurements. In this way, both models yield different but potentially complimentary models for how suicidal desire and intent evolve and vary at different timescales. We fit a bivariate hierarchical CT-VAR model with the ctsem R package (43) and a fixed-effects CT-Markov model desire and intent separately with the msm package (42). In the intensity matrix estimated by the CT-Markov model, we allow only direct transitions between adjacent states, meaning that transition intensity parameters are estimated from the zero to low, low to mild, and mild to severe states (and vice versa), but not between the zero and mild, low and severe, or zero and severe states. Descriptive analysis shows direct transitions between nonadjacent states are rarely observed in the empirical data, representing 196 (1.7%) and 213 (1.9%) out of 11,406 observed transitions for desire and intent respectively, with these transitions modeled as occurring indirectly through unobserved changes between adjacent states in the CT Markov model. As a secondary analysis, we examined whether fitting both sets of CT models separately on high-frequency (pairs of observations spaced < 1.5 h apart) and low-frequency (pairs spaced >1.5 h apart) subsets of the data (see Fig. 1 panels *B* and *C*) yielded different conclusions about the underlying dynamics. This allows us to assess whether high-frequency measurements contained unique information about the dynamics underlying suicidal thinking when compared to more typical low-frequency EMA measurements.

Data, Materials, and Software Availability. Code and scripts data have been deposited in [Github] (https://github.com/ryanoisin/TimescaleSuicidalThinking) (44). Some study data available (The data which study conclusions are based are not publicly available because the authors do not have IRB permission to publicly post the data. The deidentified data are available from the corresponding author).

ACKNOWLEDGMENTS. This research was supported by the Pershing Square Venture Fund for Research on the Foundations of Human Behavior, the Sydney DeYoung Foundation, and the Knox Fund at Harvard University. D.D.L.C. is supported by the NSF Graduate Research Fellowship under Grant No. DGE-1745303 and the National Institute of Mental Health under Grant F31MH130055. O.R. is supported by a Consolidator grant (grant agreement no. 865468) from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013). This content is solely the responsibility of the authors and does not necessarily represent the official views of the NSF or the National Institute of Mental Health. This work was also supported by the Fuss Family Research Fund at Harvard University to M.K.N.

1. S. Fazel, B. Runeson, Suicide. *N. Engl. J. Med.* **382**, 266–274 (2020).
2. J. C. Franklin *et al.*, Risk factors for suicidal thoughts and behaviors: A meta-analysis of 50 years of research. *Psychol. Bull.* **143**, 187–232 (2017).
3. K. R. Fox *et al.*, Interventions for suicide and self-injury: A meta-analysis of randomized controlled trials across nearly 50 years of research. *Psychol Bull* **146**, 1117–1145 (2020), 10.1037/bul0000305.
4. A. J. Millner, D. J. Robinaugh, M. K. Nock, Advancing the understanding of suicide: The need for formal theory and rigorous descriptive research. *Trends Cognit. Sci.* **24**, 704–716 (2020).
5. G. Turecki *et al.*, Suicide and suicide risk. *Nat. Rev. Dis. Primers* **5**, 1–22 (2019).
6. C. J. Bryan *et al.*, Nonlinear change processes and the emergence of suicidal behavior: A conceptual model based on the fluid vulnerability theory of suicide. *New Ideas Psychol.* **57**, 100758 (2020).
7. D. De Beurs *et al.*, A network perspective on suicidal behavior: Understanding suicidality as a complex system. *Suicide Life Threat. Behav.* **51**, 115–126 (2020).
8. E. M. Kleiman, M. K. Nock, Real-time assessment of suicidal thoughts and behaviors. *Curr. Opin. Psychol.* **22**, 33–37 (2018).
9. B. L. Gee, J. Han, H. Benassi, P. J. Batterham, Suicidal thoughts, suicidal behaviours and self-harm in daily life: A systematic review of ecological momentary assessment studies. *Digit. Health* **6**, 2055207620963958 (2020).
10. K. E. Adolph, S. R. Robinson, J. W. Young, F. Gill-Alvarez, What is the shape of developmental change? *Psychol. Rev.* **115**, 527–543 (2008).
11. C. J. R. Sewall, A. G. C. Wright, Personalizing suicidology. *Crisis* **42**, 405–410 (2021).
12. C. C. Driver, M. C. Voelkle, Hierarchical Bayesian continuous time dynamic modeling. *Psychol. Methods* **23**, 774–799 (2018).
13. O. Ryan, E. L. Hamaker, Time to intervene: A continuous-time approach to network analysis and centrality. *Psychometrika* **87**, 214–252 (2022).
14. W. R. Ringwald, A. G. C. Wright, Overcoming the confound of means and variability for measuring everyday emotion dynamics related to neuroticism (2022), 10.31234/osf.io/nxbyd (May 4, 2022).
15. P. Kuppens, N. B. Allen, L. Sheeber, Emotional inertia and psychological maladjustment. *Psychol. Sci.* **21**, 984–991 (2010).
16. E. M. Kleiman *et al.*, Examination of real-time fluctuations in suicidal ideation and its risk factors: Results from two ecological momentary assessment studies. *J. Abnormal Psychol.* **126**, 726–738 (2017).
17. C. J. Bryan, M. D. Rudd, The importance of temporal dynamics in the transition from suicidal thought to behavior. *Clinic. Psychol. Sci. Pract.* **23**, 21–25 (2016).
18. S. B. Wang *et al.*, A pilot study using frequent inpatient assessments of suicidal thinking to predict short-term postdischarge suicidal behavior. *JAMA Network Open* **4**, e210591 (2021).
19. A. J. Millner, M. D. Lee, M. K. Nock, Describing and measuring the pathway to suicide attempts: A preliminary study. *Suicide Life Threat. Behav.* **47**, 353–369 (2017).
20. I. Nahum-Shani *et al.*, Just-in-time adaptive interventions (JITAIs) in mobile health: Key components and design principles for ongoing health behavior support. *Annal. Behav. Med.* **52**, 446–462 (2018).
21. D. D. L. Coppersmith *et al.*, Just-in-time adaptive interventions for suicide prevention: Promise, challenges, and future directions. *Psychiatry* **85**, 317–333 (2022).
22. P. J. Batterham *et al.*, A systematic review and evaluation of measures for suicidal ideation and behaviors in population-based research. *Psychol. Assessment* **27**, 501–512 (2015).
23. A. J. Millner, M. D. Lee, M. K. Nock, Single-item measurement of suicidal behaviors: Validity and consequences of misclassification. *Plos One* **10**, e0141606 (2015).
24. D. D. L. Coppersmith, E. M. Kleiman, C. R. Glenn, A. J. Millner, M. K. Nock, The dynamics of social support among suicide attempters: A smartphone-based daily diary study. *Behav. Res. Therapy* **120**, 103348 (2019).
25. C. L. Davidson, M. D. Anestis, P. M. Gutierrez, Ecological momentary assessment is a neglected methodology in suicidology. *Archives of Suicide Res.* **21**, 1–11 (2017).
26. J. M. B. Haslbeck, O. Ryan, Recovering within-person dynamics from psychological time series. *Multivariat. Behav. Res.* **57**, 1–32 (2021).
27. C. Wrzus, A. B. Neubauer, Ecological momentary assessment: A meta-analysis on designs, samples, and compliance across research fields. *Assessment* **30**, 825–846 (2022).
28. D. D. L. Coppersmith, Effect of frequent assessment of suicidal thinking on its incidence and severity: High-resolution real-time monitoring study. *Br. J. Psychiat.* **220**, 41–43.
29. M. K. Law *et al.*, Does assessing suicidality frequently and repeatedly cause harm? A randomized control study. *Psychol. Assessment* **27**, 1171–1181 (2015).
30. R. M. Altman, Mixed hidden Markov models. *J. Am. Stat. Assoc.* **102**, 201–210 (2007).
31. N. C. Jacobson, S.-M. Chow, M. G. Newman, The differential time-varying effect model (DTVEM): A tool for diagnosing and modeling time lags in intensive longitudinal data. *Behav. Res.* **51**, 295–315 (2019).
32. J. M. B. Haslbeck, O. Ryan, D. J. Robinaugh, L. J. Waldorp, D. Borsboom, Modeling psychopathology: From data models to formal theories. *Psychol. Methods* **27**, 930–957 (2021), 10.1037/met0000303.
33. C. A. Deming *et al.*, Inconsistencies in self-reports of suicidal ideation and attempts across assessment methods. *Psychol Assess* 10.1037/pas0000976 (2021).
34. S. D. Vannoy *et al.*, Under reporting of suicide ideation in US army population screening: An ongoing challenge. *Suicide Life Threat. Behav.* **47**, 723–728 (2017).
35. J. E. Richards *et al.*, Understanding why patients may not report suicidal ideation at a health care visit prior to a suicide attempt: A qualitative study. *PS* **70**, 40–45 (2018).
36. M. A. Hom, I. H. Stanley, M. C. Podlogar, T. E. Joiner, "Are you having thoughts of suicide?" examining experiences with disclosing and denying suicidal ideation. *J. Clin. Psychol.* **73**, 1382–1392 (2017).
37. M. K. Nock *et al.*, Consensus statement on ethical & safety practices for conducting digital monitoring studies with people at risk of suicide and related behaviors. *Psychiatr. Res. Clin. Pract.* **3**, 57–66 (2021).
38. A. T. Beck, M. Kovacs, A. Weissman, Assessment of suicidal intention: The scale for suicide ideation. *J. Consult. Clin. Psychol.* **47**, 343–352 (1979).
39. O. Ryan, R. M. Kuiper, E. L. Hamaker, "A continuous-time approach to intensive longitudinal data: What, why, and how?" in *Continuous Time Modeling in the Behavioral and Related Sciences*, K. van Montfort, J. H. L. Oud, M. C. Voelkle, Eds. (Springer International Publishing, 2018), pp. 27–54.
40. D. G. Jordan, E. S. Winer, T. Salem, The current status of temporal network analysis for clinical science: Considerations as the paradigm shifts? *J. Clin. Psychol.* **76**, 1591–1612 (2020).
41. S. de Haan-Rietdijk *et al.*, On the use of mixed Markov models for intensive longitudinal data. *Multivariate Behav. Res.* **52**, 747–767 (2017).
42. C. Jackson, Multi-state models for panel data: The msm package for R. *J. Stat. Softw.* **38** (2011).
43. C. C. Driver, J. H. L. Oud, M. C. Voelkle, Continuous time structural equation modeling with R package ctsem. *J. Stat. Softw.* **77** (2017).
44. D. D. L. Coppersmith *et al.*, Code repository for "Locating the Timescale of Suicidal Thinking". https://github.com/ryanoisin/TimescaleSuicidalThinking. Deposited 2 September 2022.