scientific reports

Check for updates

OPEN

Revealing semantic and emotional structure of suicide notes with cognitive network science

Andreia Sofia Teixeira[1,2,3,4,9], Szymon Talaga[5,9], Trevor James Swanson[6] & Massimo Stella[7,8✉]

Understanding how people who commit suicide perceive their cognitive states and emotions represents an important open scientific challenge. We build upon cognitive network science, psycholinguistics and semantic frame theory to introduce a network representation of suicidal ideation as expressed in multiple suicide notes. By reconstructing the knowledge structure of such notes, we reveal interconnections between the ideas and emotional states of people who committed suicide through an analysis of emotional balance motivated by structural balance theory, semantic prominence and emotional profiling. Our results indicate that connections between positively- and negatively-valenced terms give rise to a degree of balance that is significantly higher than in a null model where the affective structure is randomized and in a linguistic baseline model capturing mind-wandering in absence of suicidal ideation. We show that suicide notes are affectively compartmentalized such that positive concepts tend to cluster together and dominate the overall network structure. Notably, this positive clustering diverges from perceptions of self, which are found to be dominated by negative, sad conceptual associations in analyses based on subject-verb-object relationships and emotional profiling. A key positive concept is "love", which integrates information relating the self to others and is semantically prominent across suicide notes. The emotions constituting the semantic frame of "love" combine joy and trust with anticipation and sadness, which can be linked to psychological theories of meaning-making as well as narrative psychology. Our results open new ways for understanding the structure of genuine suicide notes and may be used to inform future research on suicide prevention.

In many circumstances, the only difference between a completed suicide and a suicide attempt is slightly greater pressure applied to a trigger. In either case the importance of gaining a greater understanding of the psychological conditions surrounding such a tragic event is immediately apparent; what leads an individual to contemplate, and perhaps commit, such an act? Similarly, what sorts of thoughts and feelings does one encounter when experiencing suicidal ideation? And how might our understanding of these phenomena aid in improving prevention efforts? Although these are challenging questions, suicide notes represent one potential window into the psychology of individuals who complete suicide[1,2]. By analyzing the language and contents of suicide notes, we can gain unique insight into shared features of individual experiences and perhaps a greater understanding of the cognitive processes that accompany suicidal ideation[3].

Previous research on suicide notes has highlighted specific properties of such notes in an attempt to better understand what characteristics stand out and differentiate them from other types of texts. Some work has focused on studying the contents of suicide notes[4–6], including dominant emotional themes (such as "anger" and "love"), and key motives (e.g., the wish to die). In general, these studies have focused on answering the question: what is in a suicide note? That is, what are the contents that we most consistently observe when comparing notes from people that committed suicide? For example, Al-Mosaiwi and Johnstone[7] recently found that the vocabulary used by individuals at risk of suicide was different from those who suffered from other mental disorders related

[1]LASIGE, Departamento de Informática, Faculdade de Ciências, Universidade de Lisboa, 1749-016 Lisboa, Portugal. [2]INESC-ID, R. Alves Redol 9, 1000-029 Lisbon, Portugal. [3]Indiana Network Science Institute, Indiana University, 1001 IN-45, Bloomington, IN, USA. [4]Hospital da Luz Learning Health, Luz Saúde, Avenida Lusíada, 100, Edifício C, 1500-650 Lisbon, Portugal. [5]Robert Zajonc Institute for Social Studies, University of Warsaw, Stawki 5/7, Warsaw 00-183, Poland. [6]Department of Psychology, University of Kansas, 1415 Jayhawk Blvd, Lawrence, KS 66045, USA. [7]CogNosco Lab, Department of Computer Science, University of Exeter, Exeter EX4 4PY, UK. [8]Complex Science Consulting, Via Amilcare Foscarini 2, 73100 Lecce, Italy. [9]These authors contributed equally: Andreia Sofia Teixeira and Szymon Talaga. ✉email: massimo.stella@inbox.com

to depression and anxiety. Individuals who experienced suicidal ideation tended to utilize different vocabularies and mainly absolutist words, indicating that suicide notes have their own emotional and lexical footprints.

Sentiment analysis has been further applied to the goal of comparing how the emotional contents of suicide notes are categorized by learning algorithms versus trained clinicians[8,9], as well as whether or not such algorithms can reliably distinguish between genuine and simulated suicide notes[10]. These automated text-analysis techniques offer some powerful advantages over the standard, qualitative approaches that have commonly been applied to the study of suicide notes by clinical psychologists. For one, quantitative methods—such as those used in sentiment analysis[9,10] and emotional profiling[11,12]—allow for the use of more objective criteria and clearer operationalizations of psychological constructs (such as valence, emotional intensity, and others based on normative data). Qualitative methods, on the other hand, depend upon human judges to code and interpret texts, then compare ratings to assess the consistency of their conclusions. Thus, there may be a high degree of uncertainty and limited reliability when such techniques are used to make inferences. On the other hand, complex statistical/machine learning models often produce results that are difficult to understand and thus may not be very helpful for tasks different than prediction such as explanation[13].

In the present work we apply network science methods to the analysis of genuine suicide notes. Importantly, we show how network modeling can be used to expand the text-analytic toolbox in psychology and provide novel ways of answering complex research questions about text data[14]. In contrast with other automated approaches to text analysis[15], network models allow researchers to encode not only, e.g., word sentiment, but also the broader set of connections that each word has with its surrounding text[12]. This allows one to track not only *which* words appear more or less often in a sample of texts but also *how* they are used and in what contexts, thereby affording comparisons with other linguistic baselines[16]. Additionally, unlike typical black box models[13], network methods are fully transparent and produce results which are often much easier to interpret. By mapping the interrelationships among words in accordance with their usage, networks offer a clearer window into semantic frames which endow words with their broader meaning. This explicit mapping helps to both avoid the oversimplification of information conveyed in written texts as well as facilitate understanding. Additionally, network models open a window into the meaning of words by explicating meaningful semantic or syntactic relationships between concepts. Network representations make it possible to observe and understand the usage of words in terms of associations with other ideas, creating a "map" of knowledge that reflects the ways text authors think and can then be reconstructed by researchers[17]. Hence, network modeling represents an approach to the study of text data that can further elucidate the structure of human texts[18] and potentially reveal how concepts are perceived, organized and interconnected in the human mind[14].

The relevance of networks in understanding suicide notes.

Language guarantees an expression of people's perceptions through semantic content and emotions. Semantic frame theory[19] indicates that the meaning attributed by people to a given concept can be reconstructed by observing the relationships and conceptual associations attributed to that concept in text or speech. Words in a given semantic frame elicit different combinations of emotions, i.e. emotional profiles, which characterize the emotional content of a text.

Network science provides tools for quantifying and reconstructing both semantic frames[18–20] and emotional associations[12], serving as a framework for the quantitative identification of ways in which people perceive events and happenings[12,14,19,21,22]. In comparison to more opaque machine learning techniques, networks have the advantage of transparently representing a proxy for the associative structure of language in the human mind, within the cognitive system apt at acquiring, storing and producing language, i.e. the mental lexicon[21,23]. Supported by psycholinguistic inquiries into the mental lexicon[14,23–25], complex networks built from texts can open a window into people's mindsets[12]. Focus here is given to reconstructing the collective mindset as expressed in the last written words left by people who committed suicide. To this aim, we adopt a corpus of genuine suicide notes gathered in a previous study[9] and including 139 letters from people who committed suicide. The letters come from a collection of suicide notes from sources like newspapers, books and diaries collected by clinical psychologists mostly in the US and in Europe. The notes were written and collected over a time window spanning over 60 years and have been used also for recent machine learning approaches to automatic detection of suicide ideation[9]. On average, a letter included 120 ± 12 words and a total of over 2000 different concepts were stated in the whole corpus.

Cognitive network approach to suicidal ideation analysis.

In this manuscript we consider the content of suicide notes an observable realization of otherwise unobservable mental states and suicidal ideation of their authors[23]. In order to map out relationships between main concepts and the emergent semantics of suicide notes we reduced raw texts to two different network representations: co-occurrence (CO) and subject-verb-object (SVO) networks. Co-occurrence networks capture mostly successive relationships between adjacent concepts in a sentence while SVO networks capture syntactic links between actors and actions as identified with natural language processing and described in suicide letters. See Fig. 1 and "Materials and methods" section for details.

Unlike several well-known studies of semantic networks[26,27] based on semantic associations stored in lexical databases, in our approach networks of associations are extracted directly from raw texts as written down by individual people. As such it can be seen as an extension of map analysis, as used by Carley[28,29], enriched with: (i) link extraction and analysis based on modern natural language processing and network science metrics, (ii) additional cognitive data about affect patterns used in synergy with network structure, and (iii) linguistic benchmarks relying on recent datasets of conceptual associations (namely, free associations, see next section). These three points represent key ways in which we build upon and extend previous methods.
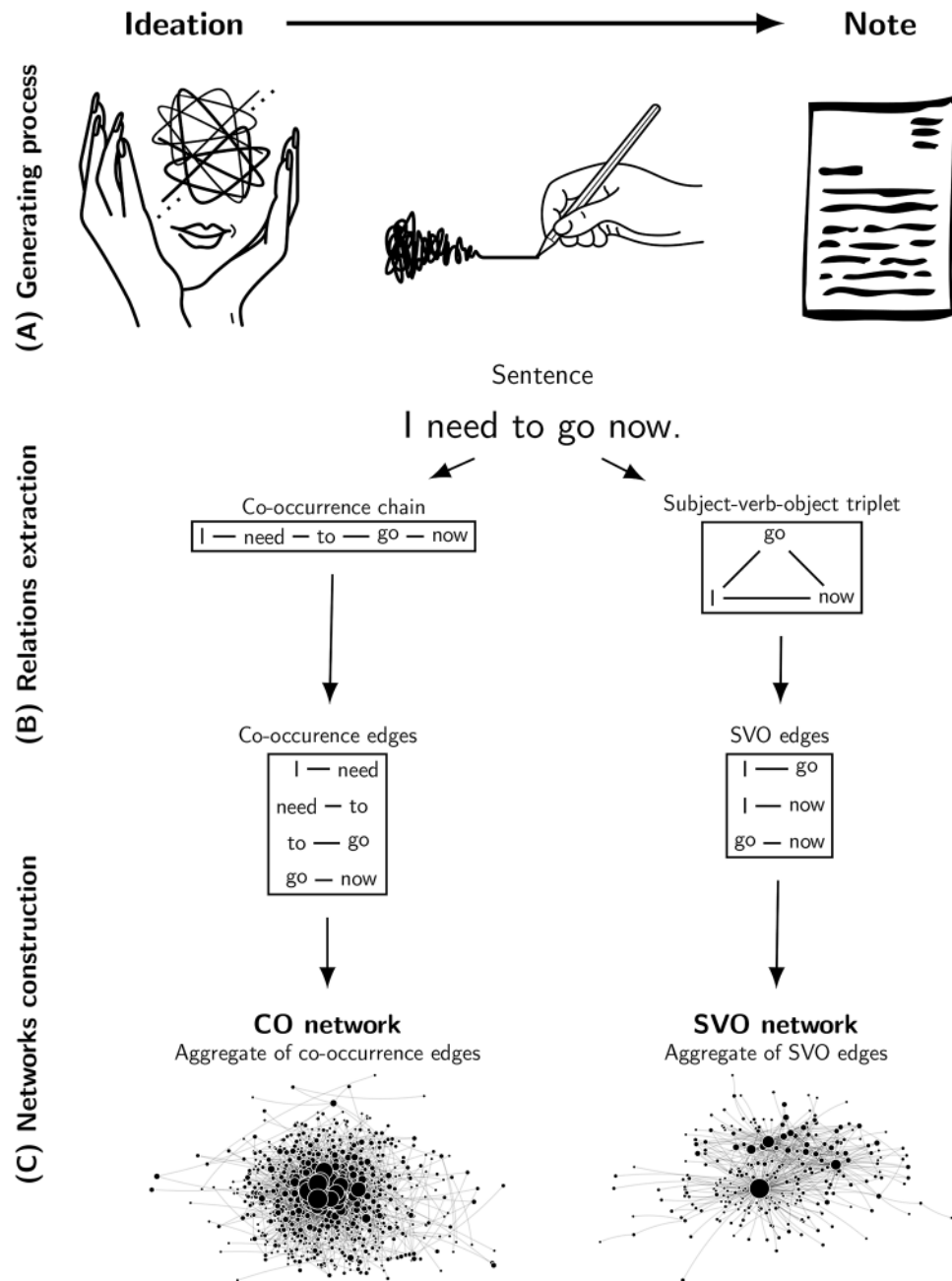
Figure 1. Schematic depiction of the data generating process (**A**), the method of relations extraction (**B**) and the network construction scheme (**C**). As (**A**) suggests, the assumption of our approach is that suicide notes are observable, even if noisy, realizations of unobservable suicidal ideation processes.

Notice that our ultimate goal is to identify mental and conceptual associations which are on average most typical for suicide notes. Hence, we do not take a sociological perspective focused on evolution of collective narrative strategies in connection with other social processes[30,31]. Instead, we aim at identifying cognitive patterns which are common across different individuals who committed suicide.

Using free associations as linguistic benchmarks for text analysis. Unlike previous approaches, we do not aim to discriminate suicide notes from other types of text[8,9]. Instead, we focus on a quantitative understanding of the mindsets of people who committed suicide. To this aim, it is important to compare the structure of conceptual associations found in suicide notes against a linguistic benchmark. Since the focus of our work lies on conceptual associations and cognitive networks, we did not select another corpus of text as a benchmark but rather utilized a separate network dataset of free associations (FA)[16,26]. These associations capture the structure of semantic memory and indicate how individuals associate concepts with one another during mind-wandering, i.e. while thinking freely without additional semantic, phonological or syntactic constraints[16]. This aspect of
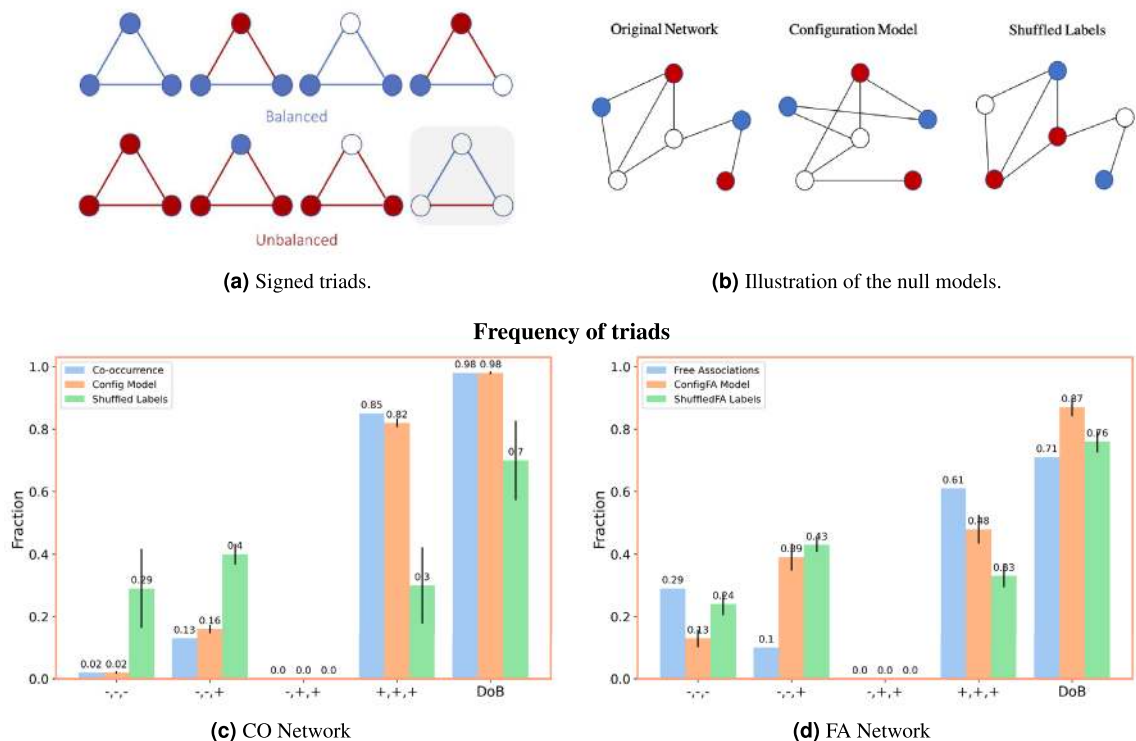
3

(a) Signed triads.

(b) Illustration of the null models.

Frequency of triads

(c) CO Network

(d) FA Network

Figure 2. On (**a**) we represent balanced and unbalanced triads along with the valence labels that generate those triads. Blue color represents positive concepts/links, red color represents negative concepts, negative links, and white circles represent neutral concepts. As explained in the text, the shadowed triad is the only unbalanced configuration impossible to obtain given any combination of positive, negative and neutral words. On (**b**) we illustrate the two null models used. Starting with a toy network we illustrate the process of rewiring links, while keeping the degree distribution with the Configuration Model, and then of shuffling the sentiment labels, while keeping the structure of the original network. On (**c**) we show the fraction of each triad and the total degree of balance for the CO Network and its correspondent null models and on (**d**) we present the same statistics for the FA network. When presenting the null models we provide the average and the standard deviation over 1000 realizations.

concept recollection makes free associations particularly appealing for investigating how the flow of specific narratives differs from mind-wandering. For this reason, we use the FA network, as obtained from the "Small-World of Words" project[16], as a linguistic benchmark for comparison against suicide notes.

Manuscript structure

Study 1 investigates the "emotional syntax" of suicide notes, analyzing whether the connectivity and configuration of words is somehow related to their valence. We use *structural balance theory*[32] to assess the degree of balance in the network and determine how valence is organized among neighboring words. We extend previous research by: (i) studying the emotional content of suicide notes and (ii) mapping how sentiment is organized in the collective mindset around suicidal ideation. Study 2 focuses on subject-verb-object relationships to highlight self-perceptions in suicide notes. Study 3 combines network centrality, semantic frames and emotional data in order to describe and quantify typical emotions associated with different concepts in suicide notes. We conclude with a general discussion of the relevance of this study *vis-á-vis* previous results and current gaps in the literature.

Study 1: investigating structural balance in suicide nodes co-occurrence networks

To assess the degree of balance of the CO network we assign signs to edges based on the presence of negative and positive sentiment labels of words (see Fig. 2a and "Materials and methods" section for more details).

Triadic closure, emotional balance and its interpretation in suicide notes. Based on sentiment labels of words, we introduce an analysis inspired by structural balance. Structural balance theory has its origin in a study by Fritz Heider, in 1946, that evaluated the psychological and cognitive configurations of interpersonal relations positioned in a triad[32]. These relations can be positive or negative and include feelings such as friendship, love, esteem, as well as their opposites. Heider stated that for a triad to be balanced it must have an even number of negative relations, otherwise tension emerges. Structural balance has also been adapted and used to study different complex systems represented as signed networks: from adaptive behavior in social networks[33,34] to financial networks[35,36], among others[37–39]. This characterization of triadic closure has also been investigated in terms of psychological states and clinical conditions[40,41]. In fact, structural balance theory has

4

been recently used to detect patterns of tension or stress in signed networks coming from cognitive neuroscience. In a study concerned with stress in the context of cooperative dilemmas, subjects engaging in unbalanced triads activated brain regions concerned with processing distress and cognitive dissonance more frequently[41]. Another recent approach shifted this psychological connotation of structural balance from social ties to the way brain regions activate or get inhibited over time in clinical populations. By analysing brain functional connectivity, Moradimanesh and colleagues showed that there were over-represented balanced triads in the brain network of people diagnosed with an autism spectrum disorder in comparison to healthy controls[40]. Notice that in the context of cognitive neuroscience, the term "balance" does not imply positive social interactions but rather positively signed brain activity (e.g. positive correlations between firings of different regions). Along this view, the increased number of balanced triads reported by[40] reflects an increased brain activity in people diagnosed with an autism spectrum disorder.

In the present study, we follow the above interpretations by extending the meaning of "balance" to encompass properties of affective organization within the context of signed networks built from conceptual associations. As such we refer to this extended interpretation as *emotional balance*. Specifically, we study how co-occurrence relationships and valences (sentiments) of words in suicide notes can provide information about possible psychological tensions, distress and emotional disturbances[41] within narrative structures. To yield positive and negative links between words, we adopt the following reasoning:

- If there is a negative word in a triad, the links connected to it will also be negative, representing emotional tension between positive and negative concepts;
- If there are two negative words, we assume the link is still negative given that the conceptual association is bridging ideas perceived as negative and thus creating additional tension;
- Links between positive words will be positive;
- If there is a neutral word in the triad, the link will retain the valence of the other end node possessing a sentiment polarity;
- If there are two or more neutral words in the triad, we do not consider the triad as being either balanced or unbalanced.

Given this setting, in Fig. 2a we show how we obtain emotionally balanced and unbalanced triads based on the valence of the words. We consider a triad emotionally balanced if there are a majority of positive words or if there is no predominant valance.

If negative concepts are predominant, we consider the triad unbalanced. In this sense, emotional unbalance reflects the preponderance of negatively-valenced words within a given triad. Our definition of emotional balance corresponds to the definition of strong structural balance considering $\{-,-,-\}$ triads as unbalanced[42].

Following this definition, a triad with a negative word will become balanced if the other two end nodes are positive. An all-positive triad emerges from all positive words or from two positive words with a neutral word. Neutral triads, since they do not contribute to a positive/negative predominance of affect, are not considered (see Fig. 2a). This emotional construction aligns with the definition of balanced ($\{+,+,+\}$ and $\{+,-,-\}$) and unbalanced ($\{-,-,-\}$) triads in structural balance[42], with the exception that in our application the formation of the unbalanced triad $\{+,+,-\}$ is not possible.

To study the emotional balance structure of mental states in suicide notes we start by building a signed network from the CO network (see "Materials and methods" section); then we evaluate its triad frequency and degree of balance—i.e., fraction of balanced triads; finally we present their statistical significance by comparing observed data against two different null models, as well as against the FA network, proceeding similarly as with the CO network. Let us underline that the comparison between the CO network, coming from suicide notes, and the FA network, coming from mind-wandering in the absence of suicide ideation, aims to identify potential patterns of psychological distress expressed in the language of suicide notes and potentially missing from the mental states of healthy controls during free mind-wandering.

Emotional balance and triad significance in suicide notes. We evaluate triad frequency in the signed CO and FA networks and evaluate their statistical significance when compared with two null models. As one model, we use configuration models[43] to generate random networks where words had the same network degree as in the CO/FA networks. We also generate random networks using the same exact structure of the CO/FA networks but shuffling the sentiment labels associated with each node. This second null model does not change the overall structure of the network but instead changes local properties (i.e. degrees) of positive, neutral and negative words relative to their characteristics in the empirical network. Notice that the signs of edges in each triad are determined by the valence of the labels associated with the nodes so different sentiment label reshufflings induce different edge signs.

We explore how triad frequency and degree of balance (DoB) in both the CO and FA networks differ from random expectation when sequential structure or affective structure is disrupted, respectively. The first comparison is obtained by configuration models randomising network structure but keeping fixed the degree of individual words. The second comparison is obtained by label-shuffling, which altered the location and local degree of words of different valence while keeping the empirical structure fixed. In Fig. 2b we present a toy example of the null models.

After creating the above modified versions of the networks, we build the associated signed networks and calculate the corresponding degree of balance and triad frequency values for both the null models. We present the mean and standard deviations for each count and for the degree of balance based on a sample of 1000 random realizations.

In Fig. 2c,d we present the results. We observe that the overall degree of balance and the frequency of $\{+, +, +\}$ triads in the CO network and its configuration models are much higher than for the label-shuffled networks (Fig. 2c). In Fig. 2d we further provide evidence that the CO network exhibits a high level of degree of balance through comparison with the FA network. In the free association network there is a more uniform distribution of triad frequencies and the null models follow the decrease of all positive triads in comparison to the co-occurrence network of suicide notes.

Emotional balance in suicide notes as a narrative strategy. The results indicate that the emotional structure of conceptual associations in suicide notes is more compartmentalized where the all-positive triads ($\{+, +, +\}$) are most frequent. Mental contents are typically described as compartmentalized when affective information is organized, with respect to a given construct, so that positive and negative associations are relatively segregated and those of the same polarity cluster together. Compartmentalized mental contents have been investigated within social psychological studies looking at vulnerability and resilience to depression[44–46]. One difference between the present usage of 'compartmentalization' is that previous researchers have focused on the affective organization of the self-concept, whereas we apply this concept to interpreting the entire mental space of authors of suicide letters. It is possible that the highly balanced pattern we observe in the co-occurrence network is indicative of a psychological strategy for coping with the "psych-ache" associated with suicidal ideation (cf.[3,4]). Namely, this higher degree of organization may suggest an active process of motivated cognition[47]. That is, the individual may be driven to associate positive concepts with one another more readily than they would in a free association context (such as mind-wandering in the absence of suicidal ideation) to increase the salience and accessibility of positive-affective information. This preliminary evidence calls for the need to conduct future survey-based research in order to investigate this possibility.

Study 2: perceptions of self via subject-verb-object relationships

In this study we use the subject-verb-object (SVO) network to identify central actors and in particular the way the self ("I") is related to other people and concepts in suicide notes. The SVO representation is more suitable for this task than the co-occurrence representation as the number of relations a given word is involved in is not determined by the number of its occurrences alone but rather by the centrality of its position in the syntactic structure of sentences in which it occurs. Hence, in the SVO representation most important sentence-building and therefore meaning-making entities have most central position in the network by design (see Network construction section in Supplementary Information (SI) for more details). Therefore, in this approach it is justified to consider node strength (sum of edge weights) an appropriate measure of centrality.

Core actors and themes. Nodes with highest strengths in clusters (Fig. 4A) as well as largest hubs in general ("I", "you", "s/he", "love", "take", "it", "go", "give", "know", "tell"; see Fig. 4C) intuitively seem to be representative of what may be typical messages that people might try to convey in their last written statements. The core actors and themes involved are self and others and relations of love (or lack thereof), taking/giving, going and telling. Figure 3 presents most frequent words and associations in the SVO network as well as the distribution of frequency of most common pairs of words among different suicide notes. Additional validation of the extracted associations is given in Table 3 (SI). Crucially, more insight into typical conceptual associations expressed in suicide notes can be gained by studying the structure of the SVO network in more detail.

As indicated by Fig. 4C the node strength distribution is heavy-tailed and hence the structure of the SVO network is determined primarily by a few hubs. Interestingly, these are personal pronouns and in particular "I", which indicates that self and other persons are main sentence-building and thus meaning-making entities in suicide notes. Clearly, the defining feature of the network is the opposition between a star-like cluster around "I" and much less centralized remaining parts of the network. Crucially, it is personal pronouns, such as "I", "you" and "s/he", which play central roles, indicating that relationships with other people (as well as self) are of crucial importance in suicide notes. This seems to reflect their personal character and focus on relations between authors and other people as well as relationships of authors with themselves.

Furthermore, the neighborhood of "I" is very star-like and its sentiment/valence is significantly more negative (19% of words) than the rest of the network (12.2%), $p = 0.040$, and less positive (26.6% relative to 36.8%), $p = 0.020$ (χ^2 test for two proportions). This suggests that perceptions and cognitions about the self are both prevalent and largely isolated from other contexts (words connected to "I" often are not linked to any other word) as well as associated with more negative emotions on average. In general, it seems that the defining feature of the network is the division between the star-like part which connects to the rest almost exclusively through "I" and the remaining part which is markedly less centralized. We make this argument more rigorous with the notion of degeneracy[50], which measures the tendency that a random walker starting from a random node after one step ends up in a limited set of central nodes (see SI: Network degeneracy) and is normalized between 0 and 1, where 1 is attained only by perfect star graphs (see Fig. 4A).

Degeneracy and component shrinkage in the SVO network. The degeneracy of the network around "I" means that there is a relatively large subset of 139 nodes (23.2%) which is linked to the giant component only through "I" (see Fig. 4B). It also means that 33.7% of negative words (29 out of 86) are linked to the network this way. No other node plays as an important role for the overall connectivity and negative sentiment. Moreover, this property is not induced by the strength distribution alone, which implies it is a higher order characteristic indicative of a special role of the self in suicidal ideation.

This phenomenon becomes evident in the comparison of the shrinkage of the giant component after node deletion in the observed SVO network relative to an appropriate null distribution based on undirected weighted
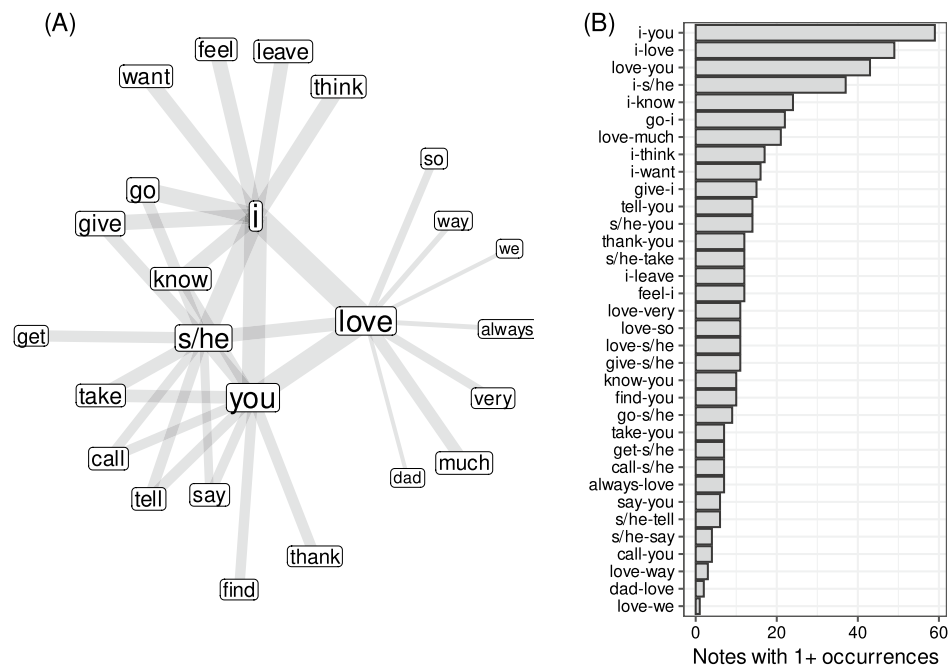
(A)

(B)

Figure 3. (A) Network of most frequent relations involving "i", "you", "s/he" and "love" derived from the SVO network. Arguably, neighbors of the above four words agree with what could be expected based the on common sense. (B) Number of documents containing specific (unidrected) pairs of concepts. Pairs which may be expected to be typical for suicide notes such as "i-you", "i-love", "love-you" and "go-i" occupy the top of the plot.

(soft) configuration model[51] (see Fig. 4B). The null model we use corresponds to a null hypothesis that words connect with each other at random conditional only on their strengths, that is, twice the number of SVO triplets they occur in. Other hubs do not differ as much in terms of their component shrinkage factors (except "s/he", but the absolute difference even in this case is much lower) and in general they tend to be either typical representatives of the null model or only slightly positive outliers. The only exception is "love" (Fig. 4B) which has significantly lower shrinkage factor, $p = 0.007$. In other words "love" has both high strength and degree and very few neighbors with no other connections. Hence, "love" is rather knitting the whole SVO network together rather than having its own specific set of associations.

SVO structure and self-perception in suicide notes.   In summary, the SVO-based analysis shows that in suicide notes: (1) the self plays a central role, in particular as the main hub through which negatively-valenced concepts are connected; (2) the large-scale structure of the concept-network is determined primarily by personal pronouns—"I", "s/he" and "you"—corresponding to self and others; (3) "love" is the main verb/noun and it does not have a significant unique neighborhood of associated concepts but rather knits together regions focused around the three main pronouns. Qualitatively, this suggests that, on average, a significant feature of suicidal ideation may be a pronounced (perhaps excessive) and largely negative focus on self as well as personal, intimate relations with other people.

This result echoes decades of research in clinical psychology[52–54] showing that negative self-schemata, as well as enhanced processing of negative self-referential information, play an important role in the emergence and persistence of depressive symptoms. Moreover, considering this finding alongside Study 1, these results are potentially suggestive of the so-called "hidden vulnerability of compartmentalization"[55]. That is, while positively-valenced mental contents may be more readily associated with one another, these are not necessarily connected with the self-concept, which may instead be dominated by negatively-valenced associations. Importantly, this interpretation may provide insight into why, in Study 1, we did not observe a prevalence of negative triads: individuals experiencing suicidal ideation may use affective compartmentalization to orient their mental associations toward positively-valenced contents in general, while concomitantly reserving negatively-valenced contents for their perceptions of themselves. Thus, negative concepts, being related primarily to the self, do not participate in many triangles in the network and this is able to explain why wee see a larger amount of $\{-, -, -\}$ triplets only in the label-reshuffled null model (see Fig. 2c).

Study 3: understanding semantic frames and emotional perceptions in suicide notes

After having investigated structural balance and self-perceptions, we turn the attention to the overall layout of conceptual and emotional perceptions in the CO network and compare these against the baseline linguistic model provided by free associations.
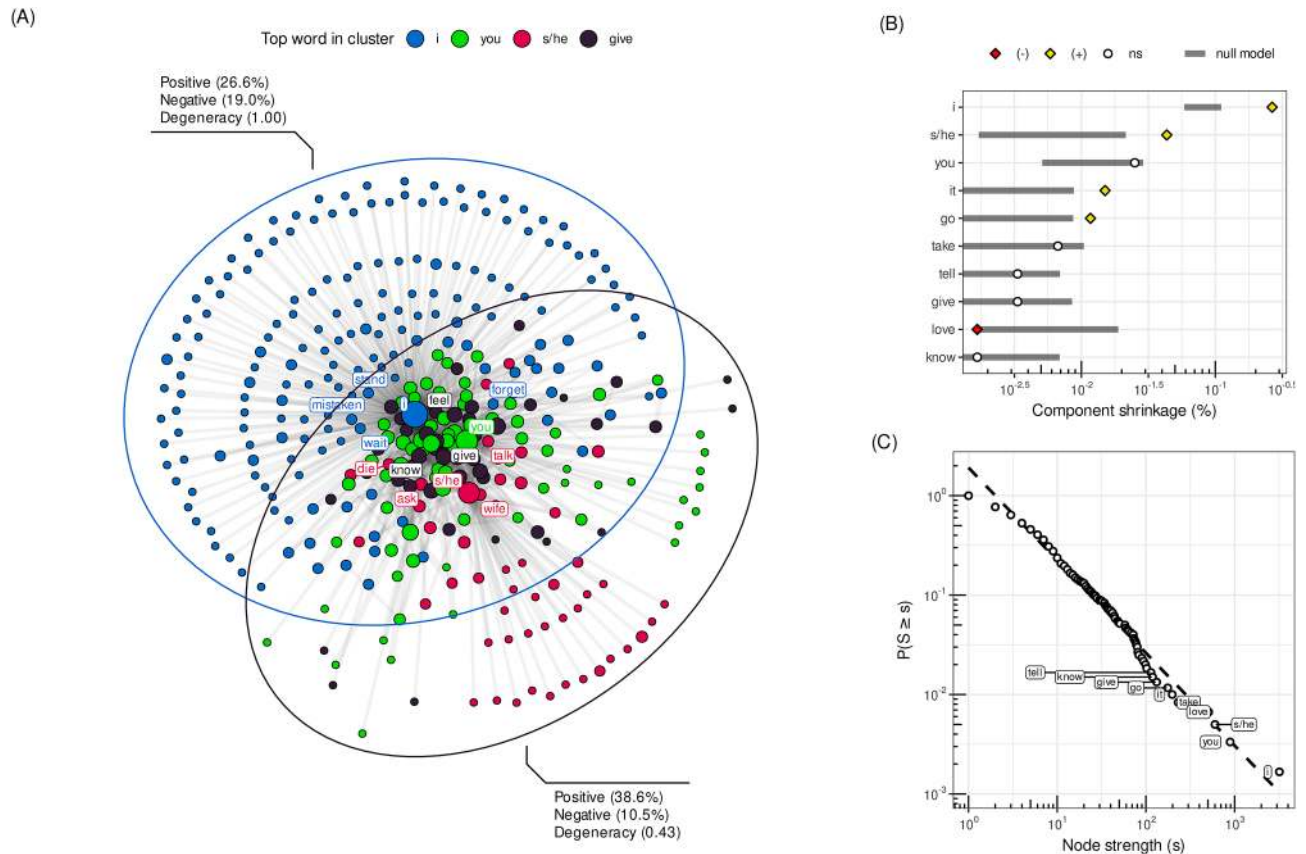
Figure 4. (**A**) Four largest clusters in the giant component of the SVO network (containing 89.7% of nodes and 98.4% of the total nodes' strength) as detected with G-N algorithm[48] (edge weights were not used). The ellipses show general orientations of the largest cluster organized around "I" and the remaining clusters. Positive/negative values refer to fractions of words with positive/negative valence. Degeneracy is between 0 and 1 and measures how a (sub)graph is similar to a star graph. (**B**) Shrinkage factors of the giant component (in log scale) after deleting a node for ten largest hubs relative to a corresponding null model distribution based on 1000 random samples (gray bounds show 1st and 99th centiles). (**C**) Complementary cumulative distribution function (CCDF) for node strength in log-log scale. The distribution is heavy-tailed (with tail exponent estimated[49] to be $\gamma \approx 2$). As a result, the topology of the network is determined primarily by a few hubs, in particular personal pronouns.

Conceptual relevance measured via network metrics. In cognitive network science, conceptual distance successfully identifies conceptual relatedness[23,24,56] (see also the SI: Semantic network distance and word prominence). The closeness ranks reported in Table 1 (SI) indicate the most prominent concepts identified in suicide notes through co-occurrence associations. Non-meaningful words such as determiners, adpositions and conjuncts were removed from the ranking as they do not convey any meaning in isolation. Among the top central words for closeness, we selected "love", "want", "help" and "life" as they were the most frequent words in the original corpus not being stopwords. We also focused on "I" in order to identify the semantic framing of the authors of suicide notes when they address themselves.

"Love" is the concept with the highest closeness centrality in the whole CO network. This indicates that suicide notes featured "love" in a wide array of different contexts and corroborates the earlier results based on the SVO representation. Other prominent concepts in the CO network include moving verbs like "go", "get", "make" and "take" and other words expressing desire such as "want", "live" and "help", which are all interconnected. "Live" and "life" are prominently featured, together with specific aspects of life such as "work" and "family". These words indicate that suicide notes are not explicitly dominated by concepts directly related to suicide, as none of the 40 top-ranked concepts reported in Table 1 (SI) are directly related to suicide or death in general. Instead, these concepts portray rather general aspects of life, even though the triad "want"/"live"/"help" indicates a willingness to get better and a focus on the topic of life.

Notice that "love" is the only top-ranked verb/noun that would significantly drop by several ranks (6) in closeness rankings based on random configuration models. This means "love" is ranked higher than random expectation in the empirical suicide notes. Free associations limited to the same set of words in common with co-occurrences identified "love" as less central than concepts like "time" and "money" (see SI: Table 2). Hence, in the process of mind-wandering free from suicidal ideation, as captured by free associations, "love" is not as

central as it is in suicide notes. These two comparisons provide compelling quantitative evidence that "love" is an exceptionally prominent concept for the authors of the notes.

The reconstruction of the semantic frame and emotions linked to "love" is thus relevant for the investigation of suicide notes. People who committed suicide might interpret or alter the commonly positive meaning attributed to "love" in mainstream language. This alteration takes place on both the semantic and emotional levels of language processing calling for additional analyses on these levels.

Emotional profiles of concepts in the CO and free association networks. Semantic prominence is not enough, on its own, to understand how authors perceived love and other concepts in their final notes. To address this, we turn to sentiment analysis and emotional profiling. Sentiment features of suicide notes helped in the supervised detection of suicidal ideation in text identification tasks (cf.[9]). In the following, emotional profiles are presented by: (i) considering the observed emotional richness of a word or set of words against random expectation; (ii) reporting the z-score of the observed richness against random word sampling from the NRC Emotion Lexicon (see SI).

Figure 5 reports the emotional profiles for highly central concepts identified in the CO network, namely "love", "want", "help", "life", and "I". The semantic frames for all these concepts include emotions that are present in suicide notes but are rather atypical for how they appear in common language (as captured by mind-wandering through free associations). This is to say that these highly central concepts are all common words but their semantic frames in suicide letters are shaped by emotions that differ drastically from the ways of thinking among healthy people associating "love", "want", "help" and "life" with the first ideas that come to their minds.

Emotions in suicide notes differ from common language. What do people who committed suicide "want"? This delicate question can be explored through the reconstructed semantic frame/emotional profile of "want". In the CO network such concept is linked with words associated with anticipation and positive emotions such as joy and trust, but also with ones eliciting negative emotions including fear and sadness. All these emotions but surprise are absent in the semantic frame of "want" coming from free associations. This difference indicates that the emotions of anticipation, joy, fear and sadness characterize "want" as a proxy for both positive and negative ideas in the context of suicidal ideation. The emotional profile and semantic frame of "want" suggest that what the authors of suicide notes express with "want" is not only the desire of positive, joyous or trustful things but also sad or frightening ideas. Differently from the common willingness to get better through positive experiences[22], the authors of suicide notes also express a willingness to get through sad or painful experiences, such as the fatal act of taking one's own life.

The concept of "help" is associated with words eliciting mostly positive emotions, featuring a high level of anticipation or projection into the future, e.g. looking for help with future events or plans. The trust associated to "help" in suicide notes is common with the linguistic baseline of free associations. In suicide notes, sadness is also featured more prominently than random expectation around "help". Notice that anticipation and sadness reveal evidence of resignation, according to Plutchik's atlas of emotions[22], an emotion absent in free mind-wandering.

"Life" is a concept with positive connotations in the free association baseline model, where it is strongly linked to trust and joy. In the language of suicide notes, the semantic frame of "life" elicits no significantly positive emotions. It is worth underlining that suicide notes frame "life" as a trust-less concept whereas in mind-wandering it has mostly trustful associations. In suicide notes, only an increased level of anticipation was found, which in itself is a rather neutral projection into the future. Both these patterns are expected in the context of final requests extending after the end of one's life[3–5,10]. Besides anticipation, "life" in suicide notes is framed in a way devoid of any positive or negative emotions, a striking result that calls for additional discussion in view of psychological theories like narrative psychology.

"Love" was found to be the most prominent concept featured in suicide notes. The semantic frame of "love" consists of words eliciting joy and trust. This indicates a positive perception of "love" itself in line with its positive perception in common language and found also in free mind-wandering, devoid of suicidal ideation, as represented here by free associations. Nonetheless, in suicide notes sadness was also present in the semantic frame of "love", as well as a reduced degree of anticipation, suggesting a more nuanced perception of this concept in comparison to free associations. Since anticipation indicates a projection of ideas into the future, the reduced anticipation and increased sadness found in suicide notes would indicate a melancholic framing of love, in contrast with the positive connotation found in mind-wandering and, thus, in common language.

In Study 2 we identified an overall negative perception of self in suicide notes. The semantic frame reconstructed in the present study better characterizes the perception of "I": self-perception in suicide notes revolves around sadness and it crucially lacks trust. This is different from the linguistic baseline represented by free associations; in fact, in the absence of suicidal ideation people framed "I" along with positive emotions like trust and joy. Both of these emotions are absent in suicide notes. This highlights a negative, trust-less perception of the self portrayed by suicide authors and it provides further evidence of psych-ache, as reported in Study 1.

The above results indicate that the meaning communicated through individual concepts is not the same as found in common language. Instead, suicide notes introduce richly nuanced contexts that shift the meaning and emotional content attributed to many concepts, making it fundamental to take these contexts into account for correctly understanding suicide notes, as pointed out by recent investigations, cf.[57].

Discussion

This first-of-its-kind study uses cognitive network science for identifying key concepts typical for suicide notes and reconstructing the meaning and emotions from the final words expressed by authors who committed suicide.
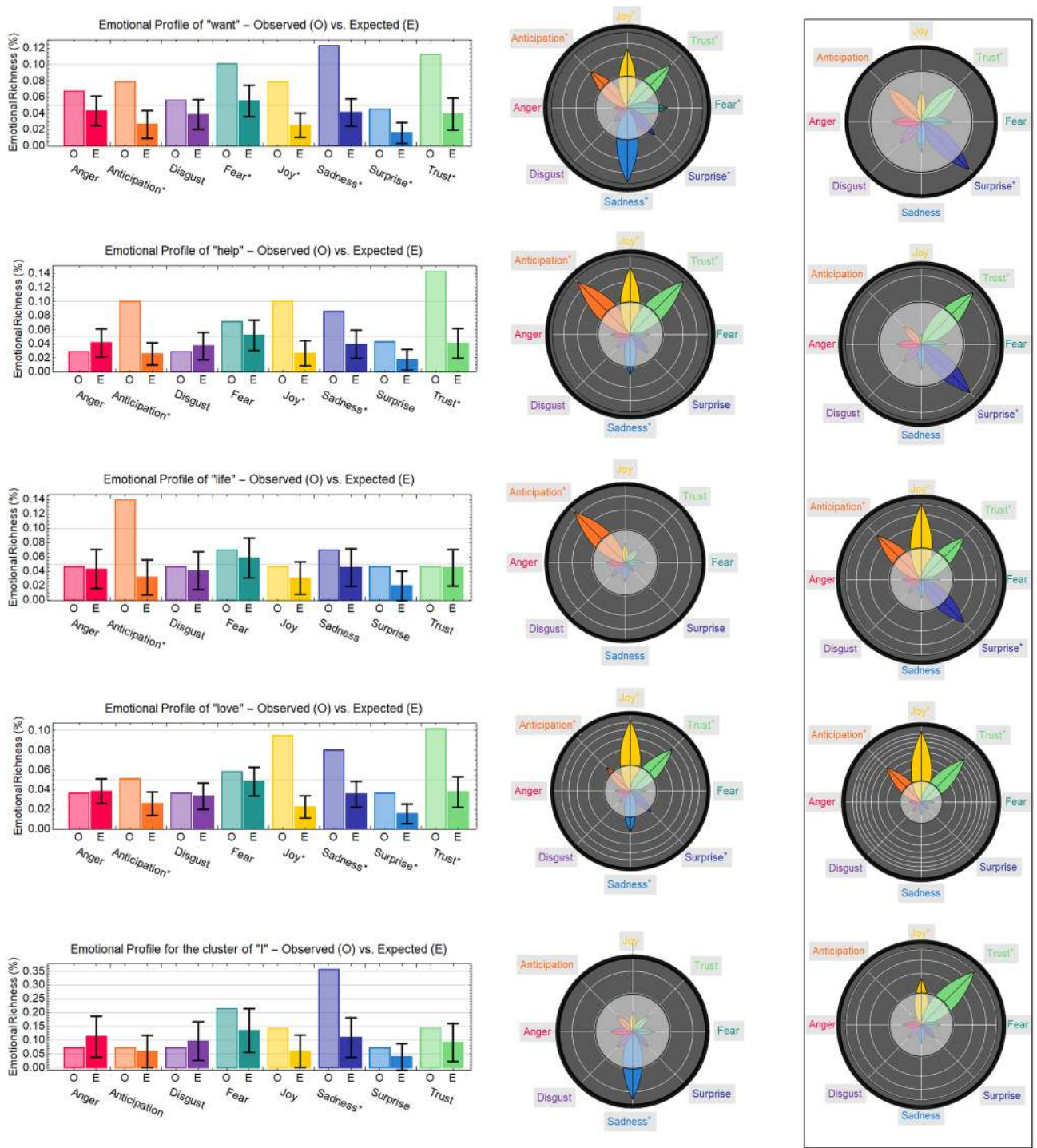
Figure 5. Reconstructed emotional profiles for "want", "help", "life", "love" and "I" in the co-occurrence network, visualized either as a comparison between observed and randomly expected emotional richness (left) or as an emotion wheel of z-scores (right). Error bars indicate one standard deviation over a sample of 1000 random iterations. Emotions more frequent than random expectation, given a significance level of $\alpha = 0.05$, are marked with an asterisk (plots on the left) and fall outside of the gray circle (plots on the right). White circles in the emotional wheels count z-scores, e.g. "life" elicited anticipation with a z-score of around 4. The framed visualizations contain emotion profiles for the same words on the left but coming from the free association network and referring to mind-wandering in the absence of suicidal ideation.

Structural balance is a potential mechanism that has been long theorized to drive certain aspects of cognitive organization, particularly with regard to resolving conflicting beliefs or attitudes[32]. Pairing this with other

psychological theories, such as *narrative psychology*[1,58] and the *meaning-maintenance model*[2,59], we can consider how the patterns observed in suicide notes fit with a broader understanding of the psychological literature.

According to the meaning-maintenance model, people are fundamentally driven to construe their lives, perceptions, and behaviors as meaningful[59]. This is a position long held by existentialist philosophers, and is widely accepted by contemporary psychologists[60]. Moreover, given the drive to make meaning from our experiences and perceptions of the world[22], people will be motivated to restore the sense of meaningfulness whenever perceptions of their own life's meaning are threatened. Suicidal ideation might represent a response to such a threat, but also poses a challenge to identifying the meaning of one's life. In this perspective, writing a suicide note may represent a way of re-establishing a sense of meaningfulness and coherence in the face of circumstances that led the individual to consider or complete suicide.

This drive to find meaningfulness in narratives, notes, and letters is supported by narrative psychology, which focuses on the function, structure, and contents of the stories we tell ourselves and others about life[1]. With meaning-making as a distal motive for writing suicide notes, we may then interpret the structural balance and positive emotional perceptions reported above as concrete signals of meaning-making, or rather as proximal communicative mechanisms through which meaning-making can be more readily achieved.

In short, we argue that: (1) people are driven to perceive their lives as meaningful and coherent, (2) they use narratives or story-telling as a way to encapsulate and restore these perceptions when threatened, and that (3) potential ways of improving the coherence of one's own psychological narratives is by introducing balance and positive emotional semantic frames to otherwise unbalanced or negative sets of cognitions.

The relevance of this reasoning to the results of the present study can be summarized in two key elements. On the one hand, our extension of structural balance[32] to networks of valenced conceptual associations provides quantitative evidence that the content of suicide notes tends to feature more positive triads than valence-reshuffled null models. This indicates that both the syntax and the valence of words used in suicide notes convey a tendency to avoid conflicting cognitions by assembling together pleasant concepts, in line with previous qualitative studies[6] and quantitative investigations using a "bag-of-words" approach[5]. On the other hand, semantic framing and emotional profiling both denoted suicide notes as being rich in positive/trustful perceptions revolving around concepts such as "love", "take", "go" and "way". These positive emotional portrayals also contained strong signals of sadness and anticipation of the future.

Notice that despite the overall prominence of positive, compartmentalized conceptual structures in the mindset of suicide notes, our analysis also found that self-perception is mostly dominated by negative associations. The semantic frame and the subject-verb-object associates of "I" highlighted negative semantic relationships of self and others, mostly dominated by sadness and absent in the linguistic baseline model provided by free associations (in absence of suicidal ideation). This represents compelling evidence for a *cognitive dissonance* in the mindset of people who committed suicide. Suicide notes denote a high level of structural balance with positive conceptual triads but also a negative cluster of concepts surrounding the self but lacking triadic closure (as evident from the SVO analysis in Fig. 4).

Our results integrate and extend previous findings and also show "love" is a central concept in suicide notes. Recent studies debated whether the emotional perception of "love" in suicide notes is as positive as in common language[57,61]. Our network approach enriched with linguistic annotations identified "love" in suicide notes as being attributed to the same positive connotations as appear in common language, but also imbued with a sense of sadness. Furthermore, while love is central across suicide notes, it is described as mostly related towards other people in diverse ways, and does not function as a purely positive emotion (as assumed by[61]). This indicates the importance of going beyond considering words in isolation[5] to better understand suicide notes. Our quantitative approach reconstructs words as interconnected in the language of suicide notes and compares it against random network models and linguistic baseline models (free associations[16]). This network structure reveals "love" as: (i) being prominent in the considered narratives, even more than in mind-wandering as captured by free associations, (ii) being focused on relations with others and (iii) eliciting a nuanced set of emotions consisting of joy and trust (as in free associations) but also nuances of anticipation and sadness. Structuring narratives around trustful and joyous relationships with loved ones aligns well with the above interpretation of suicide notes as being strategically driven by meaning identification. Another outcome of such strategy, aimed at avoiding conflicting cognitions, might be the signals of anticipation and sadness attributed to "love", which both identify resignation, a passive acceptance of threats that generates no anger or conflict at the cost of feeling defeated and incapable of creating change. This perception calls for future research investigating the psychological mechanisms at work. However, our approach already suggests a way in which emotional valences in text data such as suicide notes can be measured in a more contextualized manner by combining general population-level estimates with case-specific structure of associations between different concepts/words. Thus it can potentially be used more generally for estimating sentiment associated with different words while accounting for the specificity of a given corpus, a problem already recognized by several authors (cf.[62]).

All in all, the reconstructed contexts of concepts in suicide notes provide evidence for meaning-making narratives, aimed at coping with threats through meaning identification and conflict-avoidant storytelling. This rich landscape is invisible when considering words in isolation and only emerges from the complex structure of conceptual and emotional connections between words in text. Cognitive network science[12,14,23,24], combining psycholinguistics, computer science and network science, represents a powerful framework for reconstructing conceptual relationships, opening a window into people's minds, along with their subjective perceptions and perspectives. The ability for cognitive network analyses to parse large volumes of texts without human supervision calls for future large-scale investigations of the cognitions and perceptions expressed in suicide notes.

Limitations and future directions.   This study does not use machine learning for automatic classification of texts. It rather focuses on the reconstruction of the general mindset expressed through genuine suicide notes which represents typical streams of thoughts of people committing suicide. Achieving this quantitative knowledge is key not only for better understanding of suicide notes but also for empowering interpretable future models of automatic language processing[7–9,13].

The main limitation of the present study is the lack of comparison with a different set of texts. One might consider comparing suicide notes against other corpora, e.g. love letters[9]. However, these comparisons would include potential issues with unexpected content as found in text, as the overall topic of a corpus offers little guarantee on its linguistic content and semantic frames[15]. For instance, love letters might frame "love" in nuanced—even sad—ways because of lovestruck authors or thoughts related to regret and desire, while very similar language and mindsets can be found also in suicide letters[3,6]. Such emotional/semantic overlap may lead to considerable errors when comparing different corpora, making it difficult to produce interpretable results. In this work, we still managed to compare suicide notes against a linguistic baseline model by adopting free associations[16], which come from mind-wandering and are devoid of suicidal ideation or any other systematic frame of reference by construction. Since many of the emotional and structural differences we found between suicide notes and free associations are supported by relevant literature in clinical psychology, it suggests that our methodological framework is suitable for further analysis of autobiographical corpora, which we will address in future research.

Concerning the emotional balance of these cognitive networks, the major limitation is the inability of observing the $\{-, +, +\}$ triads, which creates a weak bias towards balanced triangles. This was only a minor issue because we were able to observe that even with this constraint, depending on the null model, different levels of balance were observed. Moreover, null models were subject to the same limitation, and so still provided appropriate comparisons within the context. A potential extension of this analysis could be based on the investigation of how to define neutral links between words. Future research might also investigate more elaborate definitions of emotional balance, based not only on valence but also on other dimensions of affect like arousal or dominance. Based on recent results in the relevant literature[63], adding multiple variables for defining balanced and unbalanced triads might reduce the observed degree of balance.

Another limitation is the identification of emotional profiles in suicide notes based on cognitive datasets referring to everyday language usage[11]. This problem is partially addressed by reconstructing emotional profiles from the specific syntactic relationships detected in suicide notes. In this way, attention should be given not to the emotions of individual words but rather to the way such words are interconnected within the observed texts. Building and adopting emotional lexical resources extracted from texts with suicidal ideation could provide more accurate readings of emotional profiles, and so represent important goals for further research.

In our approach, we operationalize the affective information of each word via norms obtained from large-scale studies grounded in affect theory[64]. It is important to note, however, that past studies have given reasons to be cautious about the fidelity of automatic operationalizations in sentiment analysis, especially due to issues of individual variability, cultural influences and training biases[65]. We note this concern, and highlight that our approach attempts to improve on this issue by reconstructing the variability of word usage in terms of network associations.

Furthermore, the present analysis is based on a corpus of a limited size and scope, both spatiotemporally and culturally. This is largely due to the limited accessibility of datasets of genuine suicide notes. Hence, our results should not be generalized without careful consideration, in particular concerning the cultural specificity of the sample. It also means that intercultural comparisons of language and mindsets expressed in suicide notes may be an interesting avenue for future research. Crucially, the approach we propose in this paper could be used to extract knowledge and cognitive structures expressed through suicide notes which vary across different populations and thus could be used to inform construction of unbiased early detection methods and systems.

Last but not least, our analysis used only „offline" letters while there is evidence that language used in online communication is often significantly different[66]. Therefore, our results should not be generalized without additional care to the context of online communication. At the same time, the general approach we propose could be in principle used to identify most prevalent differences between languages and mindsets of offline and online suicide notes.

Conclusions

In this research we present the first application of network science to the quantitative analysis of genuine suicide notes. Our approach combines some of the unique advantages of automated text analysis with networks, while using theoretical tools from psychology (e.g., structural balance theory), to gain a more detailed understanding of underlying psychological states associated with suicide notes. Cognitive network methods allow us to move beyond comparatively opaque, "black-box" models for classifying suicide notes, as they extract key ideas and emotions embedded in text. This knowledge extraction allows researchers to address questions about higher-level psychological processes and test hypotheses based on theory. Although the present study was data-driven and therefore exploratory, it demonstrates that cognitive networks are a valid approach for future confirmatory investigations, leading to a greater understanding of new possibilities for suicide prevention.

Materials and methods

Dataset of suicide notes.   A genuine suicide note is a text left by a person who subsequently committed suicide. This investigation used the Genuine Suicide Notes corpus by Schoene and Dethlefs[9]. The corpus represents a collection of 139 genuine suicide notes collected from fact-checked newspaper articles and other previous small-scale investigations of suicide notes. All notes are in English and were anonymized by changing names of people and places or any reference to identifying information. Shorter suicide notes, including those

less than two sentences, were discarded. As mentioned above, this collection of suicide notes comes from sources like newspapers, books and diaries collected by clinical psychologists mostly in the US and in Europe. Notes were written and collected over a time window spanning over 60 years, between 1958 and 2016. The corpus was assembled by Schoene and Dethlefs[9]. On average, a letter included 120 ± 12 words and a total of 2075 different concepts were stated in the whole corpus. Deleting stopwords, i.e. words not possessing a meaning when isolated from other concepts, led to 1909 different concepts being stated in the considered suicide letters.

Constructing two types of linguistic networks.

This work implemented two types of network construction. Co-occurrence (CO) networks captured syntactic relationships between adjacent words in a sentence. Subject-verb-object (SVO) networks captured triplets of syntactic relationships between a subject, a verb and an object. These network representations of the structure of knowledge in suicide nodes was also enriched with sentiment labels (e.g. a word being perceived as positive/negative/neutral in common language, see[64]) and emotional labels (i.e. words eliciting one or more emotional states, see[11]). Additional details, including a list of most common associations between concepts in the SVO networks and a set of examples of sentences including some of those associations, can be found in Supplementary Information (SI).

Emotional balance and triadic closure analysis.

Sentiment labels can be positive, negative or neutral and are used for structural balance analysis. Let $G = (V, E)$ be an undirected and signed network, with $|V|$ vertices (words) and $|E|$ edges (co-occurrence relations). We define edge labels $w \in \{-1, 1, 0\}$ between two words (a,b) as follows: $w(a, b) = w(b, a) = 0$ if both words are neutral; $w(a, b) = w(b, a) = -1$ if either a or b is negative; and $w(a, b) = w(b, a) = 1$ if both a and b are positive or if one of them is positive and the other neutral. As a result, we obtain a signed network with 1962 positive links and 1362 negative links. We consider that the neutral links (links with label 0) do not play a role when calculating the degree of balance. With the above definition the shadowed triad from Fig. 2 (unbalanced) is never observed in this signed network. The degree of balanced is obtained by calculating the fraction of balanced triads. Results were tested against configuration models[43], which randomly rewires links while keeping the graph connected and fixing the degree of each word in the empirical network. Fixing the degree of words was key in accounting for local influences of degree over non-triadic patterns like preferential attachment[67,68] and degree assortativity[26].

Free associations as a linguistic baseline model.

In the manuscript we build and investigate cognitive networks of conceptual associations representing how authors conceptually framed and perceived ideas in their last letters. These network structures have to be compared not only against baseline network null models (e.g. configuration models in emotional balance) but also against other linguistic baseline models, indicating how people in non-suicidal populations would have framed and interconnected the same set of concepts occurring in suicide notes. Whereas other investigations based on machine learning adopted textual corpora like love letters or blog posts about depression[9] as linguistic baseline models, it is not clear what type of content or semantic frames should be expected from a given text corpus. For instance, love letters might present feelings of melancholia closely related to suicidal ideation, without clear or direct control from the experimenter. This uncertainty makes the comparison more difficult. For this reason, we followed another approach, using not texts but directly complex networks as baseline linguistic models. We focused on mind-wandering, a cognitive phenomenon where concepts are interconnected with each other in ways relying on memory only and free of additional constraints (e.g. linking only concepts related syntactically within a sentence). Mind-wandering is captured by free associations[16,24], i.e. naming the first words coming up to mind when thinking of a certain concept. Hence, we use networks of free associations as baseline linguistic models representative of the mind-wandering of a large population of individuals without suicidal tendencies. Using the Small World of Words dataset[16], we built a network of free associations. It contains 1581 concepts, all included in the original co-occurrence (CO) network, which are connected according to free/mind-wandering associations. These associative structure is then used in the main text as a linguistic baseline for investigating semantic frames and emotional perceptions found in suicide notes.

Data availility
This study did not generate any new data.

References

1. McAdams, D. P. The psychology of life stories. *Rev. Gen. Psychol.* **5**, 100–122 (2001).
2. Proulx, T. & Heine, S. J. Death and black diamonds: Meaning, mortality, and the meaning maintenance model. *Psychol. Inquiry* **17**, 309–318 (2006).
3. Schneidman, E. S. Suicide notes and tragic lives. *Suicide and Life-Threatening Behavior* (1981).
4. Foster, T. Suicide note themes and suicide prevention. *Int. J. Psychiatry Med.* **33**, 323–331 (2003).
5. Handelman, L. D. & Lester, D. The content of suicide notes from attempters and completers. *Crisis J. Crisis Interv. Suicide Prev.* **28**, 102 (2007).
6. O'Connor, R. C., Sheehy, N. P. & O'Connor, D. B. A thematic analysis of suicide notes. *Crisis J. Crisis Interv. Suicide Prev.* **20**, 106 (1999).
7. Al-Mosaiwi, M. & Johnstone, T. In an absolute state: Elevated use of absolutist words is a marker specific to anxiety, depression, and suicidal ideation. *Clin. Psychol. Sci.* **6**, 529–542 (2018).
8. Pestian, J. P. *et al.* Sentiment analysis of suicide notes: A shared task. *Biomed. Inform. Insights* **5**, BII-S9042 (2012).

9. Schoene, A. M. & Dethlefs, N. Automatic identification of suicide notes from linguistic and sentiment features. In *Proceedings of the 10th SIGHUM Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities*, 128–133 (2016).

10. Pestian, J., Nasrallah, H., Matykiewicz, P., Bennett, A. & Leenaars, A. Suicide note classification using natural language processing: A content analysis. *Biomed. Inform. Insights* **3**, BII–S4706 (2010).

11. Mohammad, S. M. & Turney, P. D. Crowdsourcing a word-emotion association lexicon. *Comput. Intell.* **29**, 436–465 (2013).

12. Stella, M. Text-mining forma mentis networks reconstruct public perception of the stem gender gap in social media. arXiv preprint arXiv:2003.08835 (2020).

13. Rudin, C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat. Mach. Intell.* **1**, 206–215 (2019).

14. Castro, N. & Siew, C. S. Contributions of modern network science to the cognitive sciences: Revisiting research spirals of representation and process. *Proc. R. Soc. A* **476**, 20190825 (2020).

15. Hassani, H., Beneki, C., Unger, S., Mazinani, M. T. & Yeganegi, M. R. Text mining in big data analytics. *Big Data Cogn. Comput.* **4**, 1 (2020).

16. De Deyne, S., Navarro, D. J., Perfors, A., Brysbaert, M. & Storms, G. The, "small world of words" English word association norms for over 12,000 cue words. *Behav. Res. Methods* **51**, 987–1006 (2019).

17. Lee, M. & Martin, J. L. Coding, counting and cultural cartography. *Am. J. Cult. Sociol.* **3**, 1–33 (2015).

18. Corrêa, E. A. Jr., Marinho, V. Q. & Amancio, D. R. Semantic flow in language networks discriminates texts by genre and publication date. *Phys. A Stat. Mech. Appl.* **557**, 124895 (2020).

19. Fillmore, C. J. *et al.* Frame semantics. *Cognitive linguistics: Basic readings*, Vol. 34, 373–400 (2006).

20. Ribeiro, E., Teixeira, A. S., Ribeiro, R. & de Matos, D. M. Semantic frame induction as a community detection problem. In *International Conference on Complex Networks and Their Applications*, 274–285 (Springer, 2019).

21. Aitchison, J. *Words in the Mind: An Introduction to the Mental Lexicon* (Wiley, 2012).

22. Plutchik, R. & Kellerman, H. *Theories of Emotion* Vol. 1 (Academic Press, 2013).

23. Vitevitch, M. S. *Network Science in Cognitive Psychology* (Routledge, 2019).

24. Kenett, Y. N., Levi, E., Anaki, D. & Faust, M. The semantic distance task: Quantifying semantic distance with semantic network path length. *J. Exp. Psychol. Learn. Mem. Cogn.* **43**, 1470 (2017).

25. Stella, M., Beckage, N. M. & Brede, M. Multiplex lexical networks reveal patterns in early word acquisition in children. *Sci. Rep.* **7**, 46730 (2017).

26. Steyvers, M. & Tenenbaum, J. B. The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cogn. Sci.* **29**, 41–78 (2005).

27. Teixeira, G. M. *et al.* Complex semantic networks. *Int. J. Mod. Phys. C* **21**, 333–347 (2010).

28. Carley, K. An approach for relating social structure to cognitive structure. *J. Math. Sociol.* **12**, 137–189 (1986).

29. Carley, K. M. Extracting team mental models through textual analysis. *J. Organ. Behav. Int. J. Ind. Occup. Organ. Psychol. Behav.* **18**, 533–558 (1997).

30. Padgett, J. F., Prajda, K., Rohr, B. & Schoots, J. Political discussion and debate in narrative time: The florentine consulte e pratiche, 1376–1378. *Poetics* **78**, 101377 (2020).

31. Fuhse, J., Stuhler, O., Riebling, J. & Martin, J. L. Relating social and symbolic relations in quantitative text analysis. A study of parliamentary discourse in the Weimar Republic. *Poetics* **78**, 101363 (2020).

32. Heider, F. Attitudes and cognitive organization. *J. Psychol.* **21**, 107–112 (1946).

33. Teixeira, A. S., Santos, F. C. & Francisco, A. P. Emergence of social balance in signed networks. In *International Workshop on Complex Networks*, 185–192 (Springer, 2017).

34. He, X., Du, H., Cai, M. & Feldman, M. W. The evolution of cooperation in signed networks under the impact of structural balance. *PLoS ONE* **13**, e0205084 (2018).

35. Souto, P. C., Teixeira, A. S., Francisco, A. P. & Santos, F. C. Capturing financial volatility through simple network measures. In *International Conference on Complex Networks and their Applications*, 534–546 (Springer, 2018).

36. Harary, F., Lim, M.-H. & Wunsch, D. C. Signed graphs for portfolio analysis in risk management. *IMA J. Manag. Math.* **13**, 201–210 (2002).

37. Aref, S. & Wilson, M. C. Balance and frustration in signed networks. *J. Complex Netw.* **7**, 163–189 (2019).

38. Doreian, P. & Mrvar, A. Structural balance and signed international relations. *J. Soc. Struct.* **16**, 1 (2015).

39. Leskovec, J., Huttenlocher, D. & Kleinberg, J. Predicting positive and negative links in online social networks. In *Proceedings of the 19th international conference on World wide web*, 641–650 (2010).

40. Moradimanesh, Z., Khosrowabadi, R., Eshaghi Gordji, M. & Jafari, G. R. Altered structural balance of resting-state networks in autism. *Sci. Rep.* **11**, 1966. https://doi.org/10.1038/s41598-020-80330-0 (2021).

41. Chiang, Y.-S., Chen, Y.-W., Chuang, W.-C., Wu, C.-I. & Wu, C.-T. Triadic balance in the brain: Seeking brain evidence for Heider's structural balance theory. *Soc. Netw.* **63**, 80–90 (2020).

42. Cartwright, D. & Harary, F. Structural balance: A generalization of Heider's theory. *Psychol. Rev.* **63**, 277 (1956).

43. Viger, F. & Latapy, M. Fast generation of random connected graphs with prescribed degrees. arXiv preprint cs/0502085 (2005).

44. Showers, C. Compartmentalization of positive and negative self-knowledge: Keeping bad apples out of the bunch. *J. Personal. Soc. Psychol.* **62**, 1036 (1992).

45. Campbell, J. D., Chew, B. & Scratchley, L. S. Cognitive and emotional reactions to daily events: The effects of self-esteem and self-complexity. *J. Pers.* **59**, 473–505 (1991).

46. Showers, C. J., Abramson, L. Y. & Hogan, M. E. The dynamic self: How the content and structure of the self-concept change with mood. *J. Pers. Soc. Psychol.* **75**, 478–493 (1998).

47. Cramer, R. J., Bryson, C. N., Gardner, B. O. & Webber, W. B. Can preferences in information processing aid in understanding suicide risk among emerging adults?. *Death Stud.* **40**, 383–391 (2016).

48. Newman, M. E. & Girvan, M. Finding and evaluating community structure in networks. *Phys. Rev. E* **69**, 026113 (2004).

49. Clauset, A., Shalizi, C. R. & Newman, M. E. Power-law distributions in empirical data. *SIAM Rev.* **51**, 661–703 (2009).

50. Klein, B. & Hoel, E. The emergence of informative higher scales in complex networks. *Complexity* **1–12**, 2020. https://doi.org/10.1155/2020/8932526 (2020).

51. Squartini, T., Mastrandrea, R. & Garlaschelli, D. Unbiased sampling of network ensembles. *New J. Phys.* **17**, 023052. https://doi.org/10.1088/1367-2630/17/2/023052 (2015).

52. Beck, A. T. *Cognitive Therapy and the Emotional Disorders* (International Universities Press, 1976).

53. Ingram, R. E., Smith, T. W. & Brehm, S. S. Depression and information processing: Self-schemata and the encoding of self-referent information. *J. Pers. Soc. Psychol.* **45**, 412–420 (1983).

54. Derry, P. A. & Kuiper, N. A. Schematic processing and self-reference in clinical depression. *J. Abnorm. Psychol.* **90**, 286 (1981).

55. Zeigler-Hill, V. & Showers, C. J. Self-structure and self-esteem stability: The hidden vulnerability of compartmentalization. *Pers. Soc. Psychol. Bull.* **33**, 143–159 (2007).

56. Gray, K. *et al.* "forward flow": A new measure to quantify free thought and predict creativity. *Am. Psychol.* **74**, 539 (2019).

57. Galasinski, D. *Discourses of Men's Suicide Notes: A Qualitative Analysis* (Bloomsbury Publishing, 2017).

58. Josselson, R. & Hopkins, B. Narrative psychology and life stories. In *The Wiley Handbook of Theoretical and Philosophical Psychology* 210–233 (2015).

59. Heine, S. J., Proulx, T. & Vohs, K. D. The meaning maintenance model: On the coherence of social motivations. *Pers. Soc. Psychol. Rev.* **10**, 88–110 (2006).
60. Yalom, I. D. *Existential Psychotherapy* (Basic Books, 1980).
61. Lester, D. & Leenaars, A. A comparison of suicide notes written by men and women. *Death Stud.* **40**, 201–203 (2016).
62. Sienkiewicz, J. & Altmann, E. G. Impact of lexical and sentiment factors on the popularity of scientific papers. *R. Soc. Open Sci.* **3**, 160140 (2016).
63. Górski, P. J., Bochenina, K., Hołyst, J. A. & D'Souza, R. M. Homophily based on few attributes can impede structural balance. *Phys. Rev. Lett.* **125**, 07830 (2020).
64. Warriner, A. B., Kuperman, V. & Brysbaert, M. Norms of valence, arousal, and dominance for 13,915 English lemmas. *Behav. Res. Methods* **45**, 1191–1207 (2013).
65. Cambria, E., Poria, S., Gelbukh, A. & Thelwall, M. Sentiment analysis is a big suitcase. *IEEE Intell. Syst.* **32**, 74–80 (2017).
66. Chmiel, A. *et al.* Collective emotions online and their influence on community life. *PLoS ONE* **6**, e22207. https://doi.org/10.1371/journal.pone.0022207 (2011).
67. Hills, T. T., Maouene, J., Riordan, B. & Smith, L. B. The associative structure of language: Contextual diversity in early word learning. *J. Mem. Lang.* **63**, 259–273 (2010).
68. Hills, T. T., Maouene, M., Maouene, J., Sheya, A. & Smith, L. Longitudinal analysis of early semantic networks: Preferential attachment or preferential acquisition?. *Psychol. Sci.* **20**, 729–739 (2009).

Acknowledgements

Author contributions

A.S.T., S.T., T.J.S. and M.S. conceived and designed the experiments and interpreted the results. A.S.T., S.T. and M.S. performed the numerical experiments and contributed materials/analysis tools. All authors wrote and reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-98147-w.

Correspondence and requests for materials should be addressed to M.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.